UNIVERSITY
OF TASMANIA

Genetic control of flowering time in lentil

Vinodan Rajandran
BAgrSc(Hons)

Submitted in fulfilment of the requirements for the degree of Doctor of Philosophy

School of Biological Sciences, University of Tasmania, June 2016

Declaration of Originality

This thesis contains no material which has been accepted for a degree or diploma by the University or any other institution, except by way of background information and duly acknowledged in the thesis, and to the best of my knowledge and belief no material previously published or written by another person except where due acknowledgement is made in the text of the thesis, nor does the thesis contain any material that infringes copyright.

Authority of Access

This thesis may be made available for loan. Copying and communication of any part of this thesis is prohibited for two years from the date this statement was signed; after that time limited copying and communication is permitted in accordance with the Copyright Act 1968.

_____

Vinodan Rajandran

28th June 2016

Abstract

The transition to reproductive development is amongst the most significant in the developmental cycle of monocarpic plants. This transition, epitomised by the initiation of flowers, is of importance to agriculture. The wild progenitor of cultivated lentil, *Lens culinaris* subsp. *orientalis*, is a vernalisation responsive, facultative long day plant. The expansion of cultivated lentil, beyond the boundaries of its progenitor, coupled with the adoption of new agronomic practices, would have required the selection of landraces that were diverse in their response to both photoperiod and temperature. This diversity is reflected through the reported variation in flowering time observed in cultivated lentil. The genetic basis for this variation is of interest to plant breeding. The genetic control of flowering time in lentil however, is poorly understood. This study expands our current understanding of the genetic basis for the control of flowering time in cultivated lentil.

The current understanding of the genetic control of flowering time in lentil is shaped by the discovery of the lentil *Sn* locus. It was then proposed that the genetic control of flowering time in cultivated lentil is a function of the lentil *Sn* locus and several minor loci. The lentil *Sn* locus was first characterised in cv. Precoz, an early-flowering cultivar that has been described to be photoperiod-insensitive. This study (Chapter 3), through the use of a candidate gene approach, proposes that the *Arabidopsis ELF3* orthologue is the likely candidate for the lentil *Sn* locus, and concurs with the literature that the lentil *Sn* locus confers an early-flowering habit in its recessive state. It is also proposed that the lentil *Sn* is involved in the control of internode length, and early lateral branching. Additionally, the study suggests that the existing variation in flowering time observed in the unimproved lentil germplasm cannot be attributed to early-flowering habit conferred by the lentil *Sn*.

The *pilosae* ecotype that characterises the Indian lentil germplasm has been previously described to be early flowering. This study (Chapter 4) investigated

the genetic basis for the observed earliness in ILL 2601, a landrace that been evaluated to be amongst the earliest in the lentil germplasm. A segregating F_2 population (n=173) was established between ILL 2601 and ILL 5588 (cv. Northfield), and a genetic linkage map was constructed from 734 DArT-Seq™ markers to identify loci contributing to the observed earliness through the use of Quantitative Trait Loci (QTL) mapping. The study identified two major loci controlling flowering time, and two major loci controlling the *time to emergence from sowing*, as being integral to the observed earliness in the Indian landrace. The four described loci have not been previously reported.

This study (Chapter 5) additionally explored the genetic basis of the two major loci controlling flowering time in ILL 2601 through comparative genetics, and co-segregation analysis in segregating F_3 populations. The study proposes that the non-coding genomic sequence in the intergenic region between the *Medicago* orthologues *FTa1* and *FTa2* as having a role in conferring ILL 2601 the early-flowering habit. The study also proposes that a legume-specific paralogue of the *Arabidopsis PSEUDO-RESPONSE REGULATOR*, *PRR59c* is the likely candidate for the second locus described to confer an early-flowering habit.

Amongst the cultivated lentil germplasm, several accessions have been previously reported to be demonstrably late flowering. This study (Chapter 6) explored the genetic basis for the late flowering habit observed in cv. Indian Head. A Recombinant Inbred Line (RIL) population derived from a cross between cv. Indianhead and ILL 5588 was evaluated, and QTL determined in a genetic linkage map. Two loci contributing to the observed late flowering habit are proposed through this study.

This study, on the whole, contributes significantly through its findings to the current understanding of the genetic control of flowering time, and flowering time variation in lentil.

Acknowledgements

This thesis would not have been possible without the support, guidance and mentorship afforded by many. I would firstly like to acknowledge and thank my supervisors Jim Weller, Jules Freeman, and Valerie Hecht for their mentorship, guidance, and invaluable assistance over the past four years. I would also like to thank members past and present of my research group, in particular Lim Chee Liew, Frances Sussmilch, Jackie Vander Schoor, Stephen Ridge, and Raul Ortega-Martinez for their support and encouragement.

This research project would also not have been possible without the assistance and guidance of our collaborators. I would like to extend my appreciation to the following people; Kirstin Bett for her mentorship and guidance during my visit to the University of Saskatchewan, and access to the University of Saskatchewan Lentil Association Mapping panel (Chapter 3) and the pre-release version of LenGen (Chapter 5); Andrzej Kilian and his team at Diversity Arrays Technology Pty Ltd, Canberra for assistance with DArT genotyping (Chapter 4); and Sukhjiwan Kaur (Dimpy) and the Department of Economic Development, Jobs, Transport and Resources, Victoria, Australia for access to their cv. Indianhead x cv. Northfield RIL mapping population and assistance with linkage map construction (Chapter 6).

Closer to home, I would like to thank my parents, siblings, and my partner for their love, encouragement and support. My journey over the past few years would also not have been possible without my close friends Liam Carswell, Brendan Churchill, Jackson Tegg, Derek Sterling-Kerr, and friends from the student union movement.

I am also appreciative of the support afforded by the Endeavour Scholarships and Fellowships program and the Prime Minister's Australia-Asia Award.

Publications arising from this thesis

Journal Publications

Weller, J.L.; Liew, L.C.; Hecht, V.F.G.; Rajandran, V.; Laurie, R.E.; Ridge, S.; Wenden, B.; Vander Schoor, J.K.; Jaminon, O.; Blassiau, C.; Dalmaid, M.; Rameau, C.; Bendahmane, A.; Macknight, R.C.; Lejeune-Henaut, I. (2012). A conserved molecular basis for photoperiod adaptation in two temperate legumes. *Proceedings of the National Academy of Sciences* 109, 21158-21163.

Findings from Chapter 3 formed the basis of lentil work in Weller et al. (2012).

Conference Proceedings

Rajandran, V.; Freeman, J.; Hecht, V.; Weller, J.L. (2014). Genetic control of early flowering in Lentil. Poster presentation at VII International Conference on Legume Genetics and Genomics, Saskatoon, Canada 7-11 July 2014.

Rajandran, V.; Hecht, V.; Weller, J.L. (2012). Genetic control of flowering time in Lentil. Poster presentation at Plant Reproduction for Food 2012: XXII International Congress on Sexual Plant Reproduction, Melbourne, Australia 13-16 February 2012.

Preliminary data for Chapter 3 presented in Rajandran et al. (2012).

Preliminary data for Chapter 4 presented in Rajandran et al. (2014).

Abbreviations

AFLP	Amplified fragment length polymorphism
CAPS	Cleaved amplified polymorphic sequence
bp	Nucleotide base pairs
cDNA	Complementary DNA
CO	*CONSTANS*
COLDAIR	*COLD-ASSISTED INTRONIC NON-CODING RNA*
COLg	*CONSTANS-LIKE g*
COOLAIR	*COLD INDUCED LONG ANTISENSE INTRAGENIC RNA*
cv.	Cultivar
DFD	Delay to flower development
DNA	Deoxyribonucleic acid
DTE	Days to emergence
DTF	Days to flowering
ELF3	*EARLY FLOWERING 3*
ELF4	*EARLY FLOWERING 4*
FLC	*FLOWERING LOCUS C*
FT	*FLOWERING LOCUS T*
Gb	Giga base pair
HR	*HIGH RESPONSE*
HRM	High resolution melt
ISSR	Inter-simple sequence repeat
KASP	Kompetitive Allele Specific PCR
LATE1	*LATE BLOOMER 1*
LD	Long day
LG	Linkage group
lncRNA	Long non-coding RNA
LUX	*LUX ARRHYTHMO*
mRNA	Messenger RNA
MYB1	*V-Myb Avian Myeloblastosis Viral Oncogene Homolog*
ncRNA	Non-coding RNA
NFD	Node of flower development
NFI	Node of flower initiation
PCR	Polymerase chain reaction
PIF3c	*PHYTOCHROME-INTERACTING TRANSCRIPTION FACTOR 3c*
PRR	*PSEUDO-RESPONSE REGULATOR*
qPCR	Quantitative PCR
QTL	Quantitative trait loci
RAPD	Random amplified polymorphic DNA
SD	Short day
SE	Standard error
SN	STERILE NODES
ssp.	Subspecies
SSR	Simple sequence repeat
TOC1	*TIMING OF CAB EXPRESSION 1*

Table of Contents

List of Figures

List of Tables

List of Appendices

This page is intentionally left blank.

Chapter 1 General introduction

1.1 Introduction

The transition to reproductive development is amongst the most significant in the developmental cycle of monocarpic plants. In their respective natural environments, higher plants have evolved an elaborate mechanism that integrates both available environmental stimuli and endogenous signals to control this transition, epitomised by the initiation of flowers.

The wild progenitor of cultivated lentil, *Lens culinaris* ssp. *orientalis* (Zohary, 1972), is a vernalisation responsive, facultative long day plant. *L. culinaris* ssp. *orientalis* is endemic to the region that encompasses the Near East, the Caucasus, and Central Asia (Barulina, 1930; Zohary, 1972). The environment of this vast region is agro-ecologically diverse. This region also falls within a narrow latitudinal range (Barulina, 1930). The expansion of the cultivated form *L. culinaris* ssp. *culinaris* beyond the boundaries of its progenitor, particularly towards the equator would have imposed an intense selection pressure for local adaptation on landraces that were naturally adapted to flowering in longer days with the benefit of an extended maturation period. Equally, the adoption of new agronomic practices, including differential sowing times especially in the higher latitudes, would have imposed similar pressures. This selection for local adaptation imposed on cultivated lentil is partly reflected through the large variation in the time to flower that is described to exist within the cultivated form (Erskine et al., 1989; Erskine et al., 1994).

This thesis investigates the genetic and molecular basis for the variation in flowering time within the cultivated form.

1.2 *Lens* genus

The *Lens* Mill. genus (2n = 14) is a member of the *Fabeae* tribe that includes *Lathyrus* L., *Pisum* L., *Vavilovia* Fed., and *Vicia* L. genera. The *Fabeae* tribe includes several temperate plants of agricultural and horticultural significance, including the garden pea, common pea, grass pea, sweet pea, broad bean, and fava bean amongst others.

The current literature presents taxonomic definitions for the *Lens* genus that range from two species (Ladizinsky et al., 1984) up to six species (Mayer and Bagga, 2002). The relationship between each of the designated species within the genus is contentious, and various taxonomic definitions have been applied to both the wild and cultivated forms. While there is a growing consensus that suggests that the genus *Lens* is comprised of six species (Cubero et al., 2009; Galasso, 2003; Mayer and Bagga, 2002), namely *L. culinaris* (ssp. *culinaris* and ssp. *orientalis*), *L. odemensis*, *L. tomentosus*, *L. nigricans*, *L. ervoides*, and *L. lamottei*, more recent sequence-based assessment of the *Lens* phylogeny have suggested that some of these species may constitute a single taxon, pending further verification (Alo et al., 2011).

Wild members of the *Lens* genus are described to dominate specific geographical regions (Figure 1-1). *L. nigricans* and *L. ervoides* are primarily found in the western distribution of the genus, while *L. culinaris* ssp. *orientalis* is mainly found in the east. *L. odemensis* and *L. tomentosus* are only found in the region encompassing the Fertile Crescent and Turkey, and *L. lamottei* is specific to the region encompassing southern Europe and Morocco.

The distribution of wild members of the genus and its influence on local adaptation through hybridisation is an interesting perspective that has been put forward by several authors (Cubero, 1984; Erskine et al., 2011). Wild members of the *Lens* genus are often companion weeds of the cultivated form, and it is suggested that depending on direction of spread of *L. culinaris* spp. *culinaris*, alleles from wild members including *L. culinaris* ssp. *orientalis* and *L. odemensis*,

would have afforded local adaptation through hybridisation (Cubero, 1984; Cubero et al., 2009). Cubero et al. (2009) also argues that despite studies (Abbo and Ladizinsky, 1991; Fratini and Ruiz, 2006; Ladizinsky et al., 1985) demonstrating reproductive incompatibility and hybrid embryo abortion between the cultivated form and wild members *L. nigricans* and *L. ervoides*, the occasional successful hybrid may have contributed to the existing genetic variation within cultivated lentil.

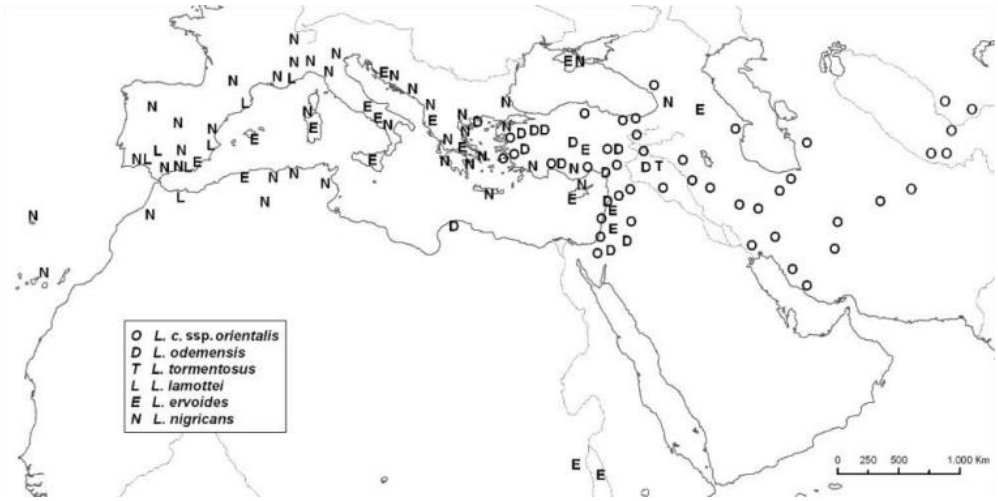


Figure 1-1 Geographical distribution of wild *Lens* species.
Distribution of wild members of *Lens* genus adapted from Cubero et al. (2009). (O) refer to members of *L. culinaris* ssp. *orientalis*, (D) refers to members of *L. odemensis*, (T) refers to members of *L. tomentosus*, (N) refers to members of *L. nigricans*, (E) refers to members of *L. ervoides*, and (L) refers to members of *L. lamottei*.

1.3 Cultivated lentil

Cultivated lentil is described as a slender, branching, softly pubescent, light green, annual plant (Barulina, 1930; Saxena, 2009). The plant is indeterminate, and displays a large variation in its growth habit, from a single stem semi-erect habit, to a vigorously branching bushy habit (Saxena, 2009). This section summaries the growth and development of cultivated lentil, previously described subgroups of cultivated lentil, and the existing genetic diversity within the crop species.

1.3.1 Stages of growth and development

The lifecycle of cultivated lentil is defined as continuous, starting from seed germination, and concluding with seed maturation (Erskine et al., 1990b). This lifecycle is divided into two distinct growth stages; the vegetative growth stage and the reproductive growth stage respectively. In lentil, the description of both vegetative and reproductive growth stages are made in relation to the identified node associated with growth and development.

Erskine et al. (1990b) designates the cotyledonary node as node 0 and progresses to define the first two nodes with simple scale-like leaves on the main stem as vegetative node V1 and V2 respectively. The first node bearing bifoliate leaves is designated V3. Erskine et al. (1990) also proposes that in the identification of vegetative growth stages in lentil, vegetative nodes are to be counted from the main stem up to the node bearing the basal primary branch, and up the basal primary branch to include the highest fully developed leaf.

The reproductive developmental stages in lentil are defined from the development of the first open flower up to pod maturity. The reproductive development stages developed and defined by Erskine et al. (1990b) are summarised in Table 1-1.

	Stage	Description
R1	First bloom	One open flower at any node.
R2	Full bloom	Open flower on nodes 10-13 of the basal primary branch.
R3	Early pod	Pod on nodes 10-13 of the basal branch.
R4	Flat pod	Pod on nodes 10-13 of the basal primary branch has reached full length and is largely flat. Seed fill is not more than half of the pod area but can be felt as a bump between the fingers.
R5	Full seed	Seed in any single pod on nodes 10-13 of the basal primary branch are swollen and pod cavity is completely filled.
R6	Full pod cavities	All the normal pods on nodes 10-13 of the basal primary branch are swollen and pod cavities are completely filled.
R7	Physiological maturity	The leaves are starting to yellow, and 50% of the pods are yellow.
R8	Full maturity	90% of the pods on the plant are golden-brown.

Table 1-1 Reproductive development stages in lentil.
Adapted from Erskine et al. (1990). Reproductive stages R1 to R8 define the reproductive development of lentil. In a community of plants, each stage is only applicable when 50% of the plants evaluated are at or have passed the designated reproductive stage.

A pre-emergent phase has also been described in the literature (Roberts et al., 1986). This phase is defined as the interval between sowing and appearance of the first shoot above the soil surface, and is described as photoperiod-insensitive (Roberts et al., 1986). The literature by and large does not account for this phase as a separate developmental phase when describing the growth and development of lentils. Moreover in previous studies (Erskine et al., 1990a; Saha et al., 2013; Tullu et al., 2008), the period for this phase is often included in the reported flowering time, recorded as the time interval from sowing to R1, associated with 50% of plants in a community in flower from sowing (Erskine et al., 1990b).

1.3.2 Subgroups of cultivated lentil

Several subgroups (ecotypes, geographical groups, races, or subspecies) of cultivated lentil are described in the existing literature. Barulina (1930) pioneered the classification of the various subgroups of cultivated lentil, and her work continues to form the basis for our current understanding of the diversity that exist within cultivated lentil.

Barulina (1930) first suggested that the cultivated form be further categorised into two distinct subgroups (Figure 1-2); namely the *microsperma* and the *macrosperma*. These classifications were largely based on pod and seed morphology, with *microsperma* accessions described for their small swollen pods and small seeds (3-6mm), and *macrosperma* accessions for their large flattened pods and large seeds (6-8mm).

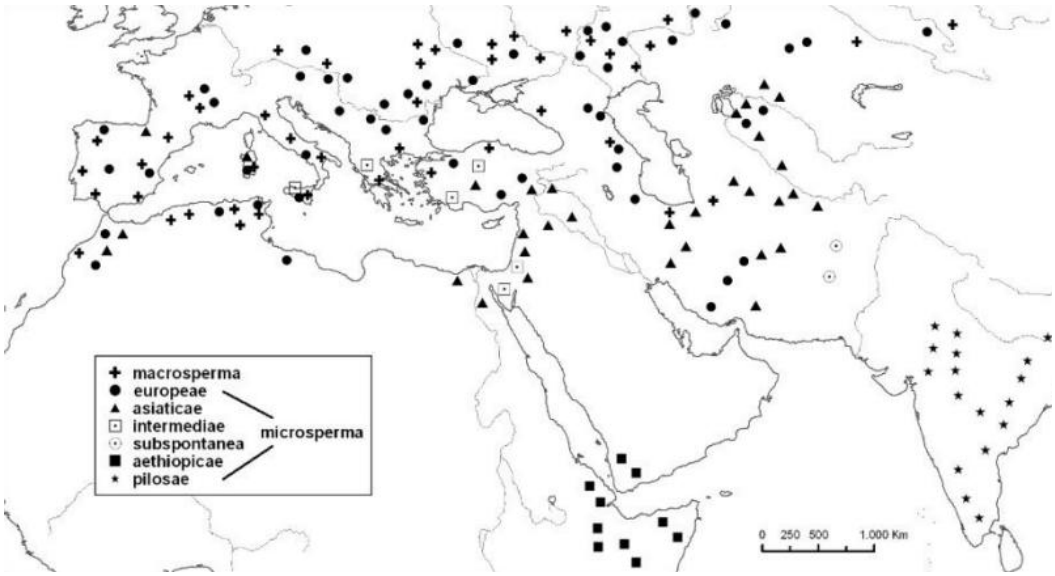


Figure 1-2 Distribution of *microsperma* and *macrosperma* cultivated lentil forms.
Adapted from Cubero et al. (2009). The distribution of both *microsperma* and *macrosperma* forms, and the six *greges* of the *microsperma* described by Barulina (1930) are illustrated.

Barulina (1930) further classified the *microsperma* accessions into six *greges*, based on their form and geographic distribution (Figure 1-2). These included the *europeae, asiaticae, intermediae, subspontanea, aethiopicae,* and *pilosae*. The latter two ecotypes are of particular interest to the study of flowering time, with both demonstrating an early-flowering phenology and reduced photoperiod sensitivity (Erskine et al., 1989; Erskine et al., 1990a; Erskine et al., 1994). Both the *aethiopicae* and *pilosae* ecotypes are geographically distributed in the lower latitudes, beyond the distribution of *L. culinaris* spp. *orientalis*, and are found in regions of high endemism (Cubero et al., 2009).

The *aethiopicae* ecotype is centred on lentils from the Ethiopian highlands, and in the neighbouring Arabian peninsular (Barulina, 1930). Cubero et al. (2009) describes the geographic region of the *aethiopicae* ecotype to overlap with the natural distribution of the wild member *L. ervoides*. Cubero et al. (2009) suggests that accessions from this region share morphological similarities with *L. ervoides*, and alludes to the likely gene flow between the *aethiopicae* ecotype and *L. ervoides*.

The *pilosae* ecotype is centred on lentils from the Indian subcontinent. This ecotype is named for its endemic traits that include pubescence on vegetative organs (Barulina, 1930), and an absence of tendrils (Vandenberg and Slinkard, 1989). Cubero et al. (2009) additionally notes that the geographic region of the *pilosae* ecotype does not overlap with the natural distribution of wild members of the *Lens* genus.

The early flowering phenology of the *pilosae* ecotype is explored in detail in Chapters 4 and 5 of this thesis.

1.3.3 Genetic diversity

Cultivated lentil is described to have a narrow genetic base (Ahmad et al., 1996; Ford et al., 1997; Sharma et al., 1995). In a study analysing both interspecific and intraspecific hybrids using Random Amplified Polymorphic DNA (RAPD) markers, Ahmad et al. (1996) determined that cultivated lentil has a narrower genetic base than most other members of the *Lens* genus, with the exception of the reproductively isolated *L. ervoides*, and *L. culinaris* ssp. *orientalis.*

Two distinct subgroups are proposed to exist within cultivated lentil (Alo et al., 2011). Alo et al. (2011) in a recent study analysing sequence alignments of introns from 22 conserved genes identified in *Medicago truncatula* proposes that the first subgroup *culinaris-M* is broadly distributed in the Mediterranean, and northern and eastern Africa, while the second group *culinaris-m* is distributed in the east, and parts of Europe (Figure 1-3). A mixture of both *culinaris-M* and *culinaris-m* was found in the Fertile Crescent. This sequence-based grouping suggests that the *pilosae* and *aethiopicae* ecotypes are from different genetic pools, and hence providing an interesting perspective on the adaptive evolution to an early flowering phenology.

Amongst the cultivated form, accessions from the Mediterranean are described to have the most intraspecific diversity (Lombardi et al., 2014), and accessions

belonging to the *pilosae* ecotype are characterised by their narrow genetic base and genetic bottleneck (Erskine et al., 2011; Ferguson et al., 1998).

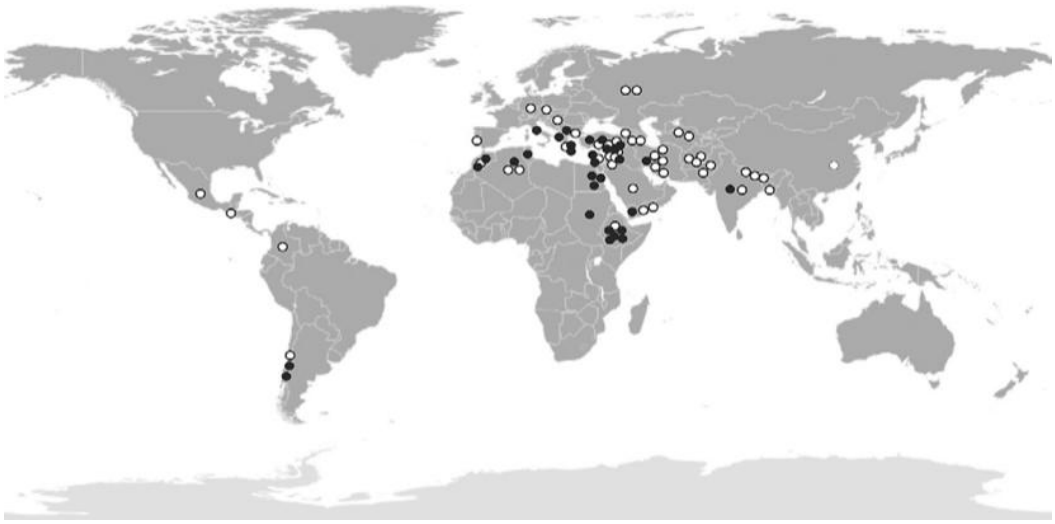


Figure 1-3 Distribution of two groups of cultivated lentil.
Distribution of *culinaris-M* (closed) and *culinaris-m* (open) subgroups (adapted from Alo et al. (2011) within cultivated lentil as determined by alignments of sequenced intronic regions of 22 genes.

1.4 Domestication and spread of cultivated lentil

1.4.1 Relationship between lentils and humans

Wild lentil seeds have been recovered in Neolithic farming settlements in the Fertile Crescent dating back to at least 7550 B.C. (van Zeist, 1970). This points to a long existing relationship between lentils and the human diet. The advent of settled societies saw the cultivation and subsequent domestication of many crop plants, including the lentil (Sonnante et al., 2009).

Lentils are amongst the earliest of crop plants that were cultivated by societies in the Fertile Crescent, together with other plants that form the Near East crop complex including wheat, barley, and pea (Abbo et al., 2009; Fuller et al., 2011a). While little is known regarding the exact period during which cultivated lentil arose from its wild progenitor *L. culinaris* spp. *orientalis*, archaeobotanical and diversity studies propose that cultivated lentil first appeared in the region encompassing modern south-eastern Turkey and northern Syria (Cubero et al., 2009; Sonnante et al., 2009). It is nonetheless, well documented that both wild

and cultivated lentil formed part of the diet of settled societies that existed in the Fertile Crescent, implying an ancient relationship between the cultivation of lentil and settled agrarian societies (Cubero et al., 2009; van Zeist, 1970).

1.4.2 Lentil domestication

The emergence of the cultivated form from the wild progenitor is a type of plant-animal co-evolutionary process (Purugganan and Fuller, 2009). In cereals, the domestication of the wild progenitor was the consequence of an intense selection by Neolithic farmers for key traits that complemented the agricultural techniques practiced (Purugganan and Fuller, 2009), and the subconscious selection for plants that were adapted to the local agro-ecological environments (Zohary, 2004), similar to that observed with natural selection. It is proposed that in cereals, traits that influenced the successful germination of wild plants under disturbed soil conditions and increased burial depth, together with traits that increased the ease of harvesting were the primary drivers of this co-evolutionary process (Purugganan and Fuller, 2009). The primary phenotypic characteristics that arose in cereals from this selection process included increased seed size and seedling vigour, germination rate, determinate growth and indehiscence (Harlan et al., 1973).

In lentils, indehiscence, conferred by a single recessive allele, *pi*, is described to have been key in the initial domestication of the wild progenitor (Ladizinsky, 1979). This 'single-step event' towards domestication is said to have offered Neolithic farmers the ease of harvesting and is proposed to have been consequently selected and maintained in the newly domesticated self-pollinating plant (Ladizinsky, 1979). It is also proposed that, together with indehiscence, a secondary trait that offered the benefits of rapid germination through reduced seed dormancy in lentils, would have been selected for by Neolithic farmers during early domestication (Ladizinsky, 1985). In *L. culinaris* ssp. *orientalis*, reduced seed dormancy conferred by a soft seed coat is controlled by a single recessive allele (Ladizinsky, 1985). The genetic basis for seed size in lentil, and its role in lentil domestication, is poorly understood

(Sonnante et al., 2009). Nonetheless, comparative analysis of archaeobotanical data suggests that the width of lentil seeds has increased in length over time as a consequence of domestication (Fuller et al., 2011a; Zohary and Hopf, 1973).

While the conscious effort to select for agronomic traits that complemented Neolithic agricultural practices dominated the co-evolutionary process during early domestication, the subconscious selection of plants that were adapted to the local agro-ecological environments influenced the rate and extent of the spread of these practices out of the Fertile Crescent (Purugganan and Fuller, 2009; Roux et al., 2006; Zohary, 2004). The adaptive evolution of these domesticated crop plants subsequently gave rise to the landraces that are now cultivated beyond the natural distribution of the wild progenitor.

1.4.3 Spread of cultivated lentil post-domestication

From the Fertile Crescent, the lentil culture followed the spread of Neolithic agricultural practices westwards to Greece, Southern Europe and North Africa, eastwards across the Middle East to central Asia and the Indian Subcontinent, and southwards to the Nile Delta and the Ethiopian highlands (Sonnante et al., 2009). The failure of both the lentil cultigen and its progenitor to adapt to the northern latitudes, unlike most other crops that form the Near East crop complex, is characteristic of its geographic spread (Colledge et al., 2005).

The expansion of the cultivated form beyond the narrow latitudinal range of *L. culinaris* ssp. *orientalis*, particularly towards the equator would have imposed an intense selection pressure upon landraces that were naturally adapted to flowering in longer days. Incidentally, of the three regions where cultivated lentil is found to be endemic (Cubero et al., 2009), two occur in regions that correspond to the distribution of the *aethiopicae* and *pilosae* ecotypes. The third centre of endemism is found in the highly diversified agro-ecological region that includes Turkmenistan, Afghanistan and Northern India (Cubero et al., 2009).

The introduction of cultivated lentil into Indo-Gangetic plains is well discussed in the literature, and is of particular interest to both plant breeding and the study of flowering time variation (Erskine et al., 1998; Erskine et al., 2011; Ferguson et al., 1998). Lentil landraces adapted to this region are entirely of the *pilosae* ecotype. Two hypotheses have been put forward (Erskine et al., 2011). The first hypothesis proposes that cultivated lentils were introduced into the Indo-Gangetic plains through Afghanistan. This was initially suggested based on linguistic evidence to have occurred with the Indo-European invasion of the Indian subcontinent (de Candolle, 1882). More recent linguistic and archaeobotanical evidence however point to an earlier land-based introduction through Afghanistan (Fuller, 2007). The alternate 'coastal route' hypothesis (Erskine et al., 2011) proposes an introduction through maritime trade. This is based largely upon similarities between the flowering phenology of the *pilosae* and *aethiopicae* ecotypes (Erskine et al., 1989), and the precedence set by the introduction of other crops such as sorghum into India which is proposed to have occurred as a consequence of trade between Africa and India (Fuller et al., 2011b).

Work characterising the diversity of the cultivated form has supported the land-based hypothesis (Ferguson et al., 1998). Through analysis isozyme and RAPD data, Ferguson et al. (1998) had determined that the *pilosae* ecotype is most similar to landraces from Afghanistan, and accordingly assigned the *pilosae* ecotype to the subgroup that included landraces from Afghanistan despite their contrasting flowering phenology. The proposed land-based introduction of lentils imply that the *pilosae* ecotype is a product of two independent selection events; first to the diversified high altitudinal region of central Asia and second to the lower latitudes of the Indian subcontinent. Both regions represent two of the three regions of endemism described for cultivated lentil (Cubero et al., 2009). In cereals, selection during domestication is estimated to result in 30-40% genome wide reduction in diversity, with loci of significance demonstrating greater reduction in diversity (Liu and Burke, 2006). By inference, the land-based hypothesis point to a significant reduction in genome wide diversity for

lentils adapted to the Indian subcontinent, consistent with the reported low genetic diversity (Ferguson et al., 1998) and described genetic bottleneck (Erskine et al., 2011) of the *pilosae* ecotype.

1.4.4 Flowering time and adaptation

Flowering time is a quantitative trait that is a function of multiple genes (Nakamichi, 2011; Salomé et al., 2011; Shrestha et al., 2014; Weller and Ortega-Martinez, 2015). In *Arabidopsis*, the adaptation of accessions to the local agro-ecological environment is facilitated by mutations at loci that have major fitness effects and others at loci with minor effects (Salomé et al., 2011).

The adaptive evolution of crop plants to diversified agro-ecological environments is described to be the consequence of an unconscious selection for local adaptation (Zohary, 2004). This process is predicted to have occurred over a short period of time unlike the protracted adaptive evolution of plants under natural selection, resulting in a preference for mutations that confer large phenotypic effects with the lowest undesirable pleiotropic effects, over mutations that confer valuable but small shifts towards the optimal phenotype (Roux et al., 2006; Zohary, 2004). This has been previously demonstrated in major crop plants including barley (Faure et al., 2012), rice (Takahashi and Shimamoto, 2011; Yano et al., 2000), and pea (Weller et al., 2012), where the selection of a single locus conferring a major-effect has been suggested to be responsible for the geographic expansion, and the adaptation of a crop to diverse agro-ecological environments. In pea for example, the *HIGH RESPONSE* (*HR*) locus, an *Arabidopsis ELF3* orthologue, is responsible for the shift to spring sowing, and the expansion of the crop to the higher latitudes (Weller et al., 2012). In barley, another *Arabidopsis ELF3* orthologue, *EARLY MATURITY 8* (*EAM8*), is proposed to have conferred adaptation to short season environments, and similarly facilitated the geographic expansion of barley cultivation (Faure et al., 2012).

Roux et al. (2006) further suggests that due to the nature of spread and the consequent successive rounds of imposed selection post-domestication, it is likely that selection of smaller-effect loci for local adaptation would have also occurred. In rice for example, this is proposed to have occurred with the independent selection of two distinct alleles of *Days to heading on Chromosome 2* (*DTH2*), an *Arabidopsis CONSTANS*-like protein, which affords minor-effect shifts for crop adaptation (Wu et al., 2013). Roux et al. (2006) has additionally proposed that fewer numbers of alleles that confer small shifts to the optimal phenotype would have been selected in the cultivated form when compared to its wild relatives.

It is suggested that the adaptive walk towards an early flowering phenotype under artificial selection is characterised by natural mutations that often result in a loss-of-function (Doebley et al., 2006; Roux et al., 2006). This is attributed to multiple pathways and rates at which loss-of-function mutants can arise. An alternative premise to the adaptive evolution of plants to an early flowering phenotype would require specific amino-acid changes that consequently result in a functional protein (Roux et al., 2006). It is also suggested that a loss-of-function mutation has the potential to confer larger phenotypic effects and therefore, a greater shift towards an optimal phenotype (Roux et al., 2006). The 'loss-of-function' pathway to early flowering has been previously observed in several crops including wheat, barley, maize and more importantly pea, a phylogenetically related legume crop (Cockram et al., 2007; Faure et al., 2012; Matsubara et al., 2012; Weller et al., 2012). A similar adaptive pathway to early flowering can be hypothesised for cultivated lentil.

1.5 Current understanding of flowering time control in lentil

1.5.1 Physiology of flowering time control in lentil

Early controlled-environment physiological studies (Erskine et al., 1990a; Roberts et al., 1988; Roberts et al., 1986; Summerfield et al., 1985) of flowering time in lentil have provided a strong basis for understanding the control and regulation of floral induction in the crop plant.

The transition to flowering in lentil is proposed to be a function of both photoperiod and temperature, with longer days and warmer temperatures evaluated to accelerate flowering (Roberts et al., 1986; Summerfield et al., 1985). It is also proposed that the relationship between photoperiod and flowering time is linear at any given temperature (Roberts et al., 1986; Summerfield et al., 1985).

The role and importance of vernalisation in floral induction for lentils however remains largely undefined. Summerfield et al (1985) in his analysis of six accessions describes a variation in vernalisation response with respect to flowering time, with vernalised plants flowering earlier in all instances when compared to non-vernalised plants. Roberts et al. (1988) in contrast proposes that that effect of vernalisation on floral induction is negligible. It has also been suggested that in vernalisation-sensitive accessions, vernalisation exposure reduces the critical or nominal base photoperiod required for floral induction (Roberts et al., 1986; Summerfield et al., 1985). More work is required to understand the role and importance of vernalisation on floral induction in lentil.

1.5.2 Diversity in lentil flowering phenology

A wide variation in flowering time is described within cultivated lentil (Erskine et al., 1989; Erskine et al., 1990a). This variation is attributed to the extent of photoperiod and temperature sensitivity of locally adapted accessions (Erskine et al., 1990a; Erskine et al., 1994). Accessions adapted to the lower latitudes are characterised by their reduced sensitivity to photoperiod and increased responsiveness to temperature, traits that collectively contribute to their early flowering phenotype (Erskine et al., 1990a; Erskine et al., 1994).

The variation in flowering phenology of cultivated lentil accessions in a global representative germplasm (Erskine et al., 1994) is illustrated together with the genetic diversity of cultivated lentil accessions from 16 countries (Ferguson et al., 1998) in Figure 1-4.

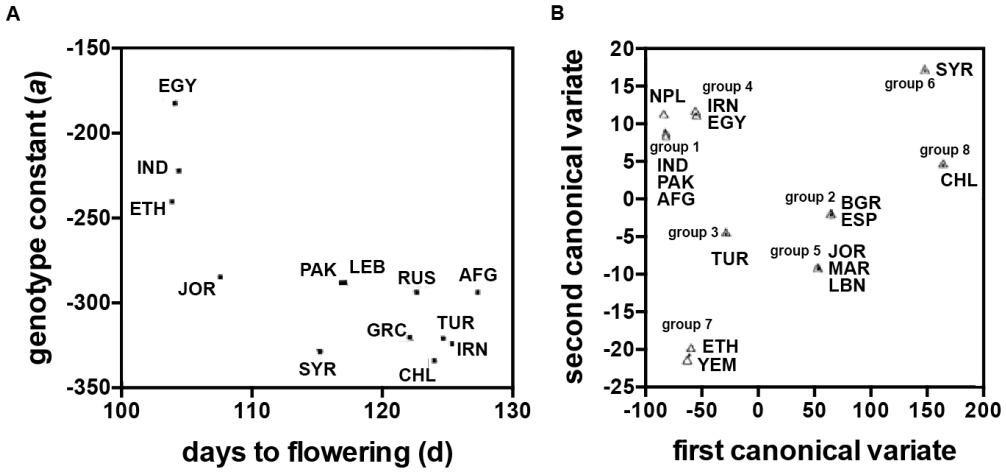


Figure 1-4 Diversity of flowering phenology in cultivated lentil.
Adapted from Erskine et al. (1994) and Ferguson et al. (1998) respectively. (A) The broad distribution of lentil flowering time in the context of their genotypic constant *a*. (B) The genetic relationship between each of the geographic subgroups with respect to their genetic diversity determined through RAPD analysis. AFG=Afghanistan, BGR=Bulgaria, CHL=Chile, EGY=Egypt, ESP=Spain, ETH=Ethiopia, GRC=Greece, IND=India, IRN=Iran, J=JOR, LEB=Lebanon, MAR=Morocco, NPL=Nepal, PAK=Pakistan, R=RUS, S=SYR, TUR=Turkey, and YEM=Yemen.

1.5.3 Genetic control of flowering time in lentil

The current understanding of the genetic control of flowering time in lentil is limited. Flowering time in lentils is proposed to be a function of a single gene at a major locus and several minor genes that confer smaller fitness effects (Sarker et al., 1999). While the current literature presents several loci (Table 1-3) involved in the control of flowering time in lentil, the lentil *Sn* locus is the only characterised flowering time locus in the crop species (Sarker et al., 1999). Furthermore, only one locus for flowering time in lentil has been mapped on a gene-based lentil genetic linkage map (Fedoruk, 2013).

The lentil *Sn* locus is demonstrated to confer an early flowering phenotype in its recessive state (Sarker et al., 1999). The recessive *Sn* allele that confers an early flowering phenotype is derived from cv. Precoz (Sarker et al., 1999), an improved accession from Argentina (Riva, 1975).

Study	Type of population	Parents	Loci	Other information
Sarker et al. (1999)	F2 Intraspecific	ILL2501 x Precoz	1	lentil *Sn,* linked to *seed coat protein* (*scp*)
	F2 Intraspecific	ILL2501 x ILL6037	1	
	F2 Intraspecific	ILL5773 x Precoz	1	
	F2 Intraspecific	ILL5773 x ILL6037	1	
Fratini et al. (2007)	RIL Interspecific	Lupa (*L. culinaris* ssp. *culinaris*) × BG16880 (*L. culinaris* ssp. *orientalis*)	3	Anonymous markers, synteny to *M. truncatula* unknown. One DTF QTL linked to *scp*, likely lentil *Sn*.
Tullu et al. (2008)	RIL Intraspecific	Eston x PI320937	2 (Saskatoon field location), 9 (Floral field location)	Anonymous markers, synteny to *M. truncatula* unknown.
Saha et al. (2013)	RIL Intraspecific	ILL6002 x ILL5588	3	Anonymous markers, synteny to M. truncatula unknown. ILL6002 is a 'pure line selection' of Precoz, one QTL is likely lentil *Sn.*
Fedoruk et al. (2013)	RIL Intraspecific	CDC Robin x 964a-46	3	Only one stable QTL. Stable QTL positioned on LG 1, syntenic to *Mt*4.0 Chr 1.
Kahriman et al. (2014)	RIL Intraspecific	Precoz x WA8649041	1	Likely lentil *Sn.*

Table 1-2 Summary of known flowering time QTL.

The characterisation of the lentil *Sn* locus suggests that the adaptive walk of the lentil cultigen to the lower latitudes post-domestication, is likely similar to that of other crop plants including barley, rice, and the garden pea, where a shift in flowering phenology is afforded by a single gene with a major fitness effect. However, unlike the examples cited in other crop plants, the prevalence of the recessive lentil *Sn* allele in the global germplasm, in particular within adapted accessions of the lower latitudes, and the genetic affinity of cv. Precoz to landraces of the Old World is not determined.

In her early work, Barulina (1930) noted that the Spanish brought in lentils cultivated in South America, and that these accessions were mainly from the *macrosperma* subgroup with limited diversity. By inference, this suggests that it is unlikely that cv. Precoz shares genetic affinity to either the *pilosae* or the *aethiopicae* ecotype. Additionally, the continuous distribution for flowering time

observed in F_2 progenies derived from crosses between cv. Precoz and its derivatives carrying the recessive lentil *Sn* allele and early flowering Indian landraces in Sarker et al. (1999) further supports this inference, suggesting a distinct genetic basis for the early phenology of the *pilosae* ecotype (Sarker et al., 1999). No *aethiopicae* accessions were included in Sarker et al. (1999).

Based on the early work by Barulina (1930) on cultivated lentil in South America, the genetic characterisation of the lentil *Sn* by Sarker et al. (1999), and diversity studies genetically separating the *pilosae* or *aethiopicae* ecotypes by Ferguson et al. (1998) and Alo et al. (2011), it can be proposed that while a convergent adaptive route to an early flowering phenology cannot be ruled out for both the *pilosae* or *aethiopicae* ecotypes, the adaptive walk towards early flowering in lentils occurred as a consequence of two or more selection events for flowering time.

1.6 Comparative understanding of genetic control of flowering time in temperate legumes

The genetic control of flowering time in temperate legumes has been explored from a comparative perspective across several legume crops (Nelson et al., 2010). In *P. sativum* (pea), the study of the genetic control of flowering time is significantly progressed (Weller et al., 2009; Weller and Ortega-Martinez, 2015). While only one locus involved in the control of flowering time has been genetically characterised in lentil and several largely anonymous loci have been described, in *P. sativum* over 20 loci (Weller and Ortega-Martinez, 2015) involved in the control of flowering time are known. This section provides a brief overview of the photoperiod and vernalisation floral induction pathways proposed in temperate legumes using examples from both *P. sativum* and *M. truncatula*.

1.6.1 Photoperiod pathway

The role of genes involved in photoperiod pathway of temperate legumes has been extensively studied and characterised (Weller and Ortega-Martinez, 2015).

Using both natural and induced mutants in *P. sativum*, the photoperiodic response in temperate legumes is characterised to be a function of both photoreceptors and genes involved in the circadian clock (Hecht et al., 2007; Liew et al., 2009a; Liew et al., 2014; Weller et al., 2004; Weller et al., 2012). Allelic variants of these genes are determined to contribute the existing variation in flowering time across various accessions of *P. sativum*. The natural occurring variants of the *P. sativum HR* locus (Weller et al., 2012) and the *STERILE NODE* (*SN*) locus (Liew et al., 2014), an *Arabidopsis LUX* orthologue, are proposed to disrupt the circadian rhythm and confer earliness to *P. sativum* accessions. The induced *P. sativum DIE NEUTRALIS* (*DNE*) mutant, an *Arabidopsis ELF4* orthologue, is likewise described to confer photoperiod-insensitivity and an early flowering phenology through the disruption of the circadian rhythm (Liew et al., 2009a). *HR*, *SN*, and *DNE* are members of the evening complex of the circadian clock in *P. sativum* (Liew et al., 2009a; Liew et al., 2014; Weller et al., 2012). Similarly, other components of the *P. sativum* circadian clock such as *LATE1* (Hecht et al., 2007), an *Arabidopsis GI* orthologue and a component of the morning complex of the circadian clock, are determined to confer a late flowering phenology through the disruption of the circadian rhythm.

Through studies in *P. sativum*, the circadian clock and its associated genes have emerged as being important to the photoperiod-dependent flowering time response of temperate legumes. In lentil a large variation in the photoperiod response from a flowering time perspective is described across a global collection of cultivated lentil. Interestingly, *Sn* was assigned to the characterised early flowering locus identified from crosses with cv. Precoz in lentil by Sarker et al. (1999) based on the early flowering phenology of the *P. sativum SN* variant described by Murfet (1971). It is likely that components of the lentil circadian clock are similarly involved in the described variation for flowering time in lentil.

1.6.2 Vernalisation pathway

The vernalisation pathway in temperate legumes is largely undefined. In *Arabidopsis*, it is proposed that a long non-coding RNA and a long antisense RNA function to regulate the *FLOWERING LOCUS C* (*FLC*) floral repressor to control the transition to reproductive development (Heo and Sung, 2011; Swiezewski et al., 2009). *FLC* orthologues are described not to exist in legumes (Hecht et al., 2005).

In *M. truncatula*, two *Arabidopsis FT* orthologues have been implicated in the vernalisation response (Laurie et al., 2011). Laurie et al. (2011) demonstrates that upon exposure to vernalising temperatures, a significant increase in the expression of both *M. truncatula FTa1* and *FTa2* is observed, with the former having a larger effect on floral induction. The molecular basis for this observation and the vernalisation response remains unknown, although it has been proposed that retroelement insertions in or 3' of *FTa1* can confer vernalisation-insensitivity (Jaudal et al., 2013). More work is required in the study of the molecular basis for the vernalisation response in legumes.

1.7 Existing genetic resources for lentil

The availability of genetic resources for the study of the lentil plant has been limited in the past. Recent advances in sequencing technologies, and a more concerted effort in developing genetic tools have made available more genetic resources for the study of the plant. This section summarises the current genetic resources available for the study of flowering time control in lentil.

1.7.1 Full genome sequencing

Advances in next generation sequencing technologies have facilitated the full genome sequencing initiative in lentil. The Lentil Genome Sequencing (LenGen) Project led by Bett et al. (2014) combines an initial draft assembly of scaffolds derived from a 23X read coverage, accounting for 2.7-Gb (approximately 60%) of the total genome, with an additional 125X read coverage of soon to be assembled scaffolds. The project is based on the Canadian accession cv. CDC Redberry (Bett et al., 2014). LenGen has recently released the v0.8 pre-release

version of the lentil genome (http://knowpulse.usask.ca/portal/lentil-genome) (Bett and Cook, 2015).

1.7.2 Transcriptome data

Transcriptome data for various lentil accessions have been made available through several published studies (Kaur et al., 2011; Sharpe et al., 2013; Temel et al., 2015; Verma et al., 2013; Vijayan et al., 2009). Prior to the availability of LenGen, published transcriptome data formed a significant resource for genetic and molecular work in lentil. A summary of publicly available transcriptome data is detailed in Table 1-3.

Study	Study description	Platform	Accessions
Vijayan et al. 2009 (unpublished)	Development of a lentil EST library. Study used mRNA extracted from young fertilized ovaries, young ovules, enlarging seeds, cotyledons of filled seed, seed coats of fully filled seeds.	Stratagene (pBluescript SK+)	Indianhead, Commando, CDC LeMay, CDC Robin, 1899T-50, 1788-4.
Verma et al. 2013	Development of a expressed gene catalogue and gene-derived functional SSR markers. Study used mRNA extracted from root and leaf tissue from plants (harvested every five days) 30-50 days after germination.	Illumina GAII	Precoz
Kaur et al. 2011	Development of molecular markers using generated transcriptome sequence data for breeding. Study used mRNA extracted from leaf (young and mature), stem, flowers, immature pods, mature pods, immature seeds, and root and shoot tiisue from seedlings.	Roche 454 Titanium	Northfield, ILL2024, ILL7537, ILL6788, Digger, Indianhead
Sharpe et al. 2013	Development of a 3'-cDNA library for a SNP Illumina GoldenGate array, and use of SNP array to construct a linkage map for a mapping population. Study used mRNA extracted from 2-week old leaf, stem before flowering, 1-week-old etiolated seedling, mixed flower stages, and developing seed at mixed stages.	Roche 454 Titanium	CDC Redberry (reference genotype), CDC Robin, CDC Milestone, Eston, 964A-46, PI 320937, LC8602303T, ILL 5588, and ILL 8006. (and two L. ervoides accessions: L01-827A and IG 72815)
Temel et al. 2015	SNP discovery and linkage map construction for Precoz x WA8649041 RIL. Study used mRNA extracted from roots, shoots, leaves, branches, and flowers for cDNA library construction.	Illumina Hiseq 2000	Precoz and WA8649041

Table 1-3 Summary of publicly available transcriptome data.

1.7.3 Molecular markers

Allele-specific molecular markers for flowering time genes in lentil were not available prior to this project.

1.7.4 Genetic linkage maps

Genetic linkage maps consisting of a range of markers have been constructed for various interspecific (Durán et al., 2004; Eujayl et al., 1998; Gupta et al., 2012b; Hamwieh et al., 2005; Tadmor et al., 1987; Tahir and Muehlbauer, 1994; Weeden et al., 1992; Zamir and Ladizinsky, 1984) and intraspecific (Gupta et al., 2012a; Kahraman et al., 2004; Kaur et al., 2014; Phan et al., 2007; Phan et al., 2006; Rubeena et al., 2003; Saha et al., 2013; Sharpe et al., 2013; Temel et al., 2015; Tullu et al., 2008) lentil populations. However, limited transferability owing to the lack of sufficient common markers across these genetic linkage maps, and the extensive use of anonymous RAPD, Inter Simple Sequence Repeats (ISSR), Amplified Fragment Length Polymorphism (AFLP), and Simple Sequence Repeat (SSR) markers have hindered a consensus on the relative position key markers.

The use of Intron-Targeted Amplified Polymorphic (ITAP) gene-based markers by Phan et al. (2006), and more recent work by Kaur et al. (2014), Sharpe et al. (2013), and Temel et al. (2015) using gene-based markers from transcriptome analysis present a new perspective on genetic linkage mapping in lentil. The genetic linkage map by Sharpe et al. (2013) is the first published intraspecific genetic linkage map constructed using gene-based markers, with seven linkage groups that correspond to the seven chromosomes of lentil.

Table 1-4 presents a summary of the lentil genetic linkage maps in the literature, from the first genetic linkage map by Zamir and Ladizinsky (1984) to the most recent by Temel et al. (2015), with details pertaining to the mapping population, total linkage map size, total number of markers, and marker type. The summary presented is not exhaustive.

Study	Type of population		Parents	No. of markers	Map length (cM)	No. of linkage groups	Type of markers
Zamir and Ladizinsky (1984)	F2	Interspecific	Four crosses between *L. culinaris ssp. culinaris* (accessions No. 2 and No. 13) and *L. culinaris ssp. orientalis* (accessions No. 22, No. 23, No. 24 ,and No. 40)	9	–	2 (six defined groups consisting of two linkage groups, rest unlinked)	8 Isozymes and one morphological marker
Tadmor et al. (1987)	F2	Interspecific	Four crosses between *L. culinaris ssp. culinaris* (accessions No. 7, No. 158, No. 160, and No. 162), *L. ervoides* (accessions No. 32), and L. odemensis (accessions No. 37)	20	258	5	18 Isozymes and two morphological markers
Weeden et al. (1992)	F2	Interspecific	No. 32 (*L. ervoides*) x No. 7 (*L. culinaris* ssp. culinaris)	66	–	11	Isozyme, RFLP, and morphological markers
Tahir and Muehlbauer (1994)	RIL	Interspecific	Eight crosses between *L. culinaris ssp. culinaris* (Brewer, Giza-9, and Redchief) and *L. culinaris ssp. orientalis* (Lo-4, Lo-56, Lo-59, Lo-66, Lo-77, and Lo-78)	21	–	6	17 Isozyme and four morphological markers
Eujayl et al. (1998)	RIL	Interspecific	ILL5588 (*L. culinaris* ssp. *culinaris*) × L692-16-1(s) (50% *L. culinaris* ssp. *orientalis*)	177	1073	7 (15 defined groups consisting of seven linkage groups and eight undefined segments)	89 RAPD, 79 AFLP, six RFLP, and three morphological markers
Rubeena et al. (2003)	F2	Intraspecific	ILL5588 x ILL7537	114	784	9	100 RAPD, 11 ISSR, and 3 RGA
Duran et al. (2004)	F2	Interspecific	Lupa (*L. culinaris* ssp. *culinaris*) × BG16880 (*L. culinaris* ssp. *orientalis*)	200	2171	10 (six with >12 markers, four with <four markers)	71 RAPD, 39 ISSR, 83 AFLP, two SSR, and five morphological markers
Kahraman et al. (2004)	RIL	Intraspecific	WA8649090 × Precoz	130	1192	9	RAPD, ISSR, and AFLP (breakdown unknown)
Hamwieh et al. (2005)	RIL	Interspecific	ILL5588 (*L. culinaris* ssp. *culinaris*) × L692-16-1(s) (50% *L. culinaris* ssp. *orientalis*)	283	751	14	39 SSR and 269 AFLP
Phan et al. (2006) and Phan et al. (2007)	RIL	Intraspecific	ILL5588 x ILL5722	97	928	7 (four unlinked markers)	79 ITAP and 18 SSR
Tullu et al. (2008)	RIL	Intraspecific	Eston × PI 320937	207	1868	12	144 AFLP, 54 RAPD, and nine SSR
Saha et al. (2013) (includes data from Saha (2009) unpublished thesis.)	RIL	Intraspecific	ILL 6002 × ILL 5888	139	1565	14	23 SSR, 27 RAPD, 89 SRAP, and two morphological markers
Gupta et al. (2012b)	RIL	Intraspecific	ILL5588 x ILL5722	196	1392	11	21 RAPD, 68 ISSR, 71 ITAP, and 36 SSR
Gupta et al. (2012a)	F2	Interspecific	L830 (L. culinaris ssp. culinaris) × ILWL77 (L. culinaris ssp. orientalis)	199	3843	11	162 RAPD, 28 SSR, and nine ISSR
Sharpe et al. (2013)	RIL	Intraspecific	CDC Robin × 964a-46	561	597	7	547 SNP, 10 SSR, and four morphological markers
Kaur et al. (2014)	RIL	Intraspecific	Cassab × ILL 2024	318	1178	10 (seven markers unlinked)	57 SSR and 261 SNP
Temel et al. 2015	RIL	Intraspecific	Precoz × WA8649041	388	433	9 (seven major groups and two minor group)	376 SNP, three SSR, and nine ISSR

Table 1-4 Summary of genetic linkage maps developed in lentil.

1.7.5 Syntenic relationship between lentil and M. truncatula

The relationship between the seven chromosomes of lentil to that of the model legume plant *M. truncatula* had been ambiguous until the recent comparative analysis presented by Sharpe et al. (2013). While previous (Phan et al., 2007;

Phan et al., 2006), and more recent (Kaur et al., 2014) publications have presented various iterations of this relationship, deficiencies in the published genetic linkage maps and associated comparative analysis have hindered a comprehensive understanding of synteny between lentil and the model legume. Sharpe et al. (2013) reports that the relationship lentil between and *M. truncatula* is largely linear, with the exception a major collinear translocation of *M. truncatula* chromosome 6 to the middle of lentil linkage group 2. A translocation of the ends of lentil linkage group 1 and 5 is also described in the context of *M. truncatula*. Several inversions and minor translocations are also described (Sharpe et al., 2013).

Chapter 4 of this thesis presents an updated iteration of this relationship.

1.8 Aims of this study

This study aims to progress the current understanding of flowering time control in lentil by investigating the genetic and molecular basis for the variation in flowering phenology within the cultivated form.

From an adaption perspective the variation in flowering phenology, in particular early flowering, is likely to have been derived from two or more independent selection events. To elucidate the genetic and molecular basis for the early flowering phenology, Chapter 3 will seek to first determine the molecular identity of the lentil *Sn* locus. This allows the study to determine the contribution of the lentil *Sn* locus to the observed variation for flowering time in the lentil germplasm, and if the lentil *Sn* was crucial for adaptation to the lower latitudes in the Old World.

Chapter 4 will then progress to genetically characterise loci contributing to the flowering phenology of the *pilosae* ecotype. The *pilosae* ecotype represents amongst the earliest flowering lentil accessions, and understanding the genetic control of the early phenology of this group of cultivated lentil will present opportunities for crop improvement. This chapter will genetically characterise an early selection of the Indian landrace ILL 2601.

Chapter 5 attempts to understand the molecular basis for the loci identified to be responsible for the early flowering phenology of ILL 2601. The chapter will also seek to understand the contribution of identified loci to existing variation for flowering time in the lentil germplasm.

Chapter 6 presents a different perspective on flowering time by seeking to understand the genetic basis for the control of the late-flowering phenology in lentil. This chapter will study cv. Indianhead, an improved accession that is amongst the latest to flower in our cultivated lentil collection. This chapter aims to provide the basis for future work in the study of the genetic control of flowering time in lentil.

Chapter 2 General materials and methods

This chapter details the general materials and methods employed in this thesis. Laboratory and product-specific protocol descriptions detailed are adapted from manufacturer's instructions. Details pertaining to chapter specific materials and methods are detailed in the materials and methods section of each chapter.

2.1 Plant materials and growth conditions

The passport information of accessions and their derivatives used for segregation analysis in this thesis are summarised in Table 2-1. Refer to chapter specific materials and methods for passport information pertaining to germplasm collections surveyed in individual chapters.

Accession	Other names	Country	Latitude	Longitude	Altitude (m)	Source	Miscellaneous information
ILL 6005	-	Argentina	-27.0	unknown	unknown	Erskine, W.	Early flowering selection from ILL 4605 (cv. Precoz) x ILL 4349 (cv. Laird) cross (Erskine, W., pers. comm.)
ILL 223	ATC71102	Iran	38.7	46.3	1360.0	Australian Temperate Field Crops Collection	
ILL 5588	Northfield	Jordan	32.1	35.8	860.0	Erskine, W.	Mid-late flowering and photoperiod-sensitive (Weller and Murfet, unpublished)
ILL 2601	-	India	23.0	unknown	unknown	Erskine, W.	Early flowering landrace (Erskine, W., pers. comm.)
Indianhead	ILL 481	Lebanon (Erskine, W., pers. comm.)	unknown	unknown	unknown	Australian Temperate Field Crops Collection	Late flowering accession (Vandenberg, A., pers. comm.)

Table 2-1 Passport information for accessions.
Summary of accessions used for segregation analysis in this thesis. Refer to Appendix for full list of accessions studied in this thesis.

In this thesis, all segregating populations were established from crosses between the accession of interest and ILL 5588 (cv. Northfield). ILL 5588 is a single-plant selection from the Jordanian landrace NEL 16, and was developed by the International Centre for Agricultural Research in Dry Areas (ICARDA) for

its resistance to *Fusarium oxysporum* and *Ascochyta lentis* (Erskine et al., 1996). ILL 5588 is described to be mid-late flowering, and is photoperiod-sensitive (Weller and Murfet, unpublished).

For all experiments in this thesis, seeds were scarified and imbibed in autoclaved Milli-Q water (Milli-Q Plus, Merck Millipore, USA) for 12-hours prior to sowing. Seeds were sown in 14 cm slim-line pots containing a 1:1 mixture of dolerite chips and vermiculite, topped with a soil media containing 1:1 mixture of sterile nursery grade potting mix with controlled release fertilizer and granulated sand. Plants were lightly watered regularly, and a nutrient solution applied weekly.

All plants described in this thesis were grown under controlled photoperiod conditions in a phytotron at the University of Tasmania. Phytotrons were maintained at approximately 24^{o}C during the day and 16^{o}C at night. Refer to individual chapter for details pertaining to photoperiod conditions for specific experiments.

2.2 Plant measurements

Details of measured plant traits are described in Table 2-2. All traits detailed in the thesis were measured as per description unless stated otherwise. All lengths were measured to nearest millimetre. Plants exhibiting abnormal growth were excluded from analysis.

Trait		Details
BTF	Branches to flower development	Number of branches to the development of the first developed/open flower on main stem. Branches are recorded when length exceeds 5 mm.
DFD	Delay to flower development	Node interval between NFI and NFD. Measured in nodes.
DTE	Days to emergence	Number of days to seedling emergence from sowing. The appearance of the first two open leaves are recorded as the time of emergence. Measured in days.
DTF	Days to flowering	Number of days to the appearance of the first developed/open flower on main stem from seedling emergence (DTE). Measured in days.
EBL	Length of early branches	Total length of branches for the first three weeks from emergence. Branches are recorded when length exceeds 5 mm. Branch lengths are measured in millimetres.
EBN	Number of early branches	Total number of branches for the first three weeks from seedling emergence (DTE). Branches are recorded when length exceeds 5 mm.
IN9	Internode interval between nodes 1 and 9	Length between nodes 1 and 9. Measured in millimetres.
IN15	Internode interval between nodes 9 and 15	Length between nodes 9 and 15. Measured in millimetres.
NFD	Node of flower development	Number of nodes to the development of the first developed/open flower on main stem. Measured in nodes.
NFI	Node of floral initiation	Number of nodes to the initiation of the first floral structure on main stem. Measured in nodes.
PH	Plant height	Total length of main stem measured from the first node. Plant heights are measured in millimetres.
PPN	Pods per node	Average number of pods per node recorded for first three reproductive nodes.

Table 2-2 Details of measured plant traits.

2.3 Online resources

Online resources for genomic and expressed sequences, and germplasm databases for passport information used in this thesis are detailed in Table 2-3.

Full genomic sequences (http://knowpulse2.usask.ca/) for lentil were only available in September 2014 (Bett, K., pers. comm.). Prior to the availability of full genome sequence, publicly available expressed sequences (http://www.ncbi.nlm.nih.gov/) were used. Passport information from both the Genesys (https://www.genesys-pgr.org/) and United States Department of Agriculture Agricultural Research Service's Germplasm Resources Information

Network (http://www.ars-grin.gov/) were used to define the geographic affinity of accessions studied in this thesis.

Resource type	Species	Website	Version
Genomic and expressed sequences	Arabidopsis (*Arabidopsis thaliana*)	http://www.arabidopsis.org/	TAIR10
	Black cottonwood *(Populus trichocarpa)*	http://phytozome.jgi.doe.gov/	v3.0
	Chickpea (*Cicer arietinum*)	http://cicar.comparative-legumes.org/	v1.0
	Common bean (*Phaseolus vulgaris*)	http://phytozome.jgi.doe.gov/	v1.0
	Lentil (*Lens culinaris*)	http://knowpulse2.usask.ca/	v0.7
		http://www.ncbi.nlm.nih.gov/	-
	Medicago (*Medicago truncatula*)	http://jcvi.org/medicago/	v4.0
	Pea (*Pisum sativum*)	http://www.ncbi.nlm.nih.gov/	-
	Purple false brome (*Brachypodium distachyon*)	http://phytozome.jgi.doe.gov/	v2.1
	Rice (*Orizya sativa*)	http://phytozome.jgi.doe.gov/	v7.0
	Soybean (*Glycine max*)	http://phytozome.jgi.doe.gov/	Wm82.a2.v1
Passport information	Lentil (*Lens culinaris*)	http://www.ars-grin.gov/	-
		https://www.genesys-pgr.org/	-

Table 2-3 Details of online resources.

2.4 Primer design

Primers were designed against available lentil sequences or in regions of sequence conservation resolved through nucleotide alignment between *M. truncatula*, *P. sativum*, and other legumes. Primers were designed using the web-based Primer3 application (http://primer3.wi.mit.edu/). Primers were optimised for primer length (18 to 24-bp), product length, G/C content, annealing temperature, minimal self or cross compatibility, and the presence of a GC clamp at 3' end. Refer to Appendix 2 and Appendix 3 for primer information.

2.5 DNA and RNA extractions and processing

2.5.1 Standard genomic DNA extraction

Genomic DNA was extracted from plant tissue samples collected in liquid nitrogen and stored at -70^{o}C prior to processing. Tissue samples were ground using a carbide bead and a mechanical tissue lyser (Qiagen TissueLyserII). Mortar and pestles were used for grinding when genomic DNA with minimal shearing was required. Ground samples were stabilised, and nucleic material extracted using 500 µl of 2x Extraction buffer (100mM Tris-HCl, 1.4M NaCl, 20mM EDTA, 2% w/v CTAB, 20mM 2-β-mercaptoethanol, pH 8 with HCl) and incubated for 15 min at 60^{o}C with gentle agitation. Solvent extraction using chloroform-isoamylic alcohol (24:1) solution was carried out twice to purify extracted nucleic material. Nucleic material was subsequently precipitated with 1mL of Precipitation Buffer (50mM Tris-HCl, 10mM EDTA, 1% w/v CTAB, pH 8 with HCl), and pelleted by centrifugation for 10 min at 14,000g. Pellets were then suspended in 300µL of 1.5M NaCl containing 1µL RNase A (25mg/mL), and incubated for 10-15 min at 50°C. Genomic DNA was precipitated in 95% ethanol, pelleted by centrifugation at 14,000g for 15 min. The genomic DNA was then washed in 70% ethanol, air dried, and dissolved in autoclaved Milli-Q water.

2.5.2 RNA extraction and cDNA synthesis

RNA was extracted from tissue samples using the Promega SV Total RNA Isolation System (Promega, USA) as described in the manufacturer's instructions. All frozen tissue samples were ground either using mortar and pestle or a carbide bead and a mechanical tissue lyser (Qiagen TissueLyserII) prior to extraction.

Complementary DNA (cDNA) strands were synthesised from 1µg RNA using the ImProm-II™ Reverse Transcription System (Promega, USA), Tetro Reverse Transcriptase (Bioline, UK), or MMLV High Performance Reverse Transcriptase (Epicentre, USA), each in a total volume of 20µL as described in the manufacturer's instructions. All samples were checked for contamination using

a negative control without reverse transcriptase. cDNA product obtained were diluted 1:5 before use.

2.6 Polymerase Chain Reactions (PCR)

2.6.1 Standard PCR

Standard PCR was carried out in 50µL volume reactions. Each reaction comprised of 5µL of template DNA, 10µL of 5x reaction buffer, 1µL of dNTPs (10mM), 1µL of forward primer (10µM), 1µL of reverse primer (10µM), 1.5µL $MgCl_2$ (50mM), and 0.2µL of MangoTaq™ DNA polymerase (Bioline, Australia), with autoclaved Milli-Q water to final volume. Reactions were performed in a thermal cycler for 30-40 cycles using the following temperature configurations: 94°C for 5 min, single cycle (94°C for 45 seconds, annealing temperature for 45 seconds, 72°C for 1 min per kb of expected product size), 72°C for 10 min.

Phusion® High-Fidelity DNA polymerase (New England BioLabs, USA) and RANGER DNA polymerase (Bioline, Australia) were used for high fidelity PCR, with reaction mix adjusted accordingly. Refer to manufacturer's instructions for more information (https://www.neb.com/products/m0530-phusion-high-fidelity-dna-polymerase and http://www.bioline.com/au/ranger-dna-polymerase.html).

2.6.2 Colony PCR

Colony PCR was carried out using bacterial colonies suspended in 5µL of autoclaved Milli-Q water in 20µL volume reactions. Bacteria were lysed by incubation at 95°C for 5 min prior to PCR. Each reaction comprised of 5µL of template DNA, 5µL of 5x reaction buffer, 0.5µL of dNTPs (10mM), 0.5µL of forward primer (10µM), 0.5µL of reverse primer (10µM), 0.8µL $MgCl_2$ (50mM), and 0.1µL of MangoTaq™ DNA polymerase (Bioline, Australia), with autoclaved Milli-Q water to final volume. Reactions were performed in a thermal cycler for 30 cycles using the following temperature configurations: 94°C for 5 min, single cycle (94°C for 1 min, annealing temperature for 1 min, 72°C for 1 min per kb of expected product size), 72°C for 5 min.

2.6.3 Quantitative PCR (qPCR)

Relative gene expression was measured using qPCR. qPCR was conducted using the Rotor-Gene Q machine (Corbett Research, Australia) operating on the Rotor-Gene 6 v6.1 software. The CAS-1200™ pipetting robot (Corbett Research, Australia), operating on the CAS Robotics v4.9.8 software, was employed to prepare reactions. Each 10µL reaction was comprised of 2µL cDNA template, 5µL 2x Quantace SensiMixPlus SYBR reagent (Bioline, Australia), 0.3µL each of forward and reverse primer (10µM), and 2.4µL autoclaved Milli-Q water. For each reaction run, a no template control was included to assess for contamination, and all samples were run in duplicate for increased result reliability. For each sample, *Elongation factor 1-α* (*Ef1-α*) was run on the reverse transcriptase negative control to check for contamination. Reactions were run for 50 cycles.

A standard curve ($R^2 \geq 0.99$) for the target gene was included in each run. Standard curves were generated from a 10-fold serial dilution from 10^{-1} to 10^{-7} ng/µL. Gene expression levels were calculated relative to *Ef1-α*, based on non-equal amplification efficiencies and the deviation in threshold cycle using the means of two technical replicates.

2.6.4 Visualisation of DNA

Amplified or digested DNA products were fractionated using electrophoresis on agarose gel in TAE buffer (40mM Tris Acetate and 1mM EDTA), stained with GoldView™ Nucleic Acid Stain (Acridine orange; SBS Genetech, China), and visualised under UV light. A DNA ladder appropriate to expected product size was used to estimate product size.

2.6.5 PCR product purification

PCR products were purified using Promega Wizard® SV Gel and PCR Clean-Up System (Promega, USA), and suspended in autoclaved Milli-Q water as per manufacturer's instructions.

2.6.6 Rapid amplification of cDNA ends (RACE)

RACE was carried out using the Clontech SMART RACE cDNA Amplification Kit (Clontech, USA) as detailed in the manufacturer's instructions. RACE was conducted with cDNA strands (section 2.5.2) synthesised from total RNA extracted from shoot and leaf tissue. Gene-specific and nested primers were designed as per manufacturer's guidelines. RACE products were visualised using gel electrophoresis.

2.7 Cloning

Cloning was carried out by ligating purified PCR products into pGEM®-T Easy vectors (Promega, USA), as per manufacturer's instructions. Vectors were subsequently inserted into competent *Escherichia coli* cells using electroporation at 1200V. Transformed cells were allowed to recover in 400µL of Luria Broth (LB, 10g/L Bacto-tryptone, 5g/L Bacto- yeast extract, 10g/L NaCl, pH 7.5) with incubation at 37°C for 1-hour with shaking. Transformed reactions were spread across LB agar (15g/L agar with 100µg/mL ampicillin and 1µL/mL X-gal) plates and incubated for 12-hours at 37°C. Transformed colonies were screened for an insert of desired length by colony PCR.

2.8 Quantification of DNA, RNA and PCR products

DNA, RNA and PCR products were quantified using a NanoDrop 8000 Spectrophotometer (Thermo Fisher Scientific, USA) as per manufacturer's instructions.

2.9 Sequencing and sequence analysis

DNA products were sequenced at Macrogen Inc. (Seoul, Korea). Sequences were edited manually using Sequencher v4.8 (Gene Codes, USA) to correct for falsely identified bases, and to remove unintelligible sequence regions at the 3' and 5' ends. Sequences were annotated using DNASTAR® Lasergene SeqBuilder v8.1.4 (DNASTAR Inc., USA).

2.10 Molecular marker design for mapping and genotyping

Molecular markers for both mapping and genotyping were designed by isolating and amplifying full or partial genomic DNA sequences of genes of interest from parental accessions of segregating populations analysed in this thesis. PCR products were purified and sequenced, and polymorphisms between parental accessions analysed for the design of an appropriate molecular marker.

2.10.1 Cleaved Amplified Polymorphic Sequence (CAPS) markers

Restriction enzyme recognition sites resulting from allelic polymorphisms between lentil parental accessions for genes of interest were used to develop CAPS markers. Primers were designed around these recognition sites to amplify DNA fragments for enzyme digests. Genomic regions that resulted in digested products having differing sizes for each parental accession that are visible during gel electrophoresis were selected for marker design. Enzyme digests were conducted as per manufacturer's instructions (New England BioLabs Inc., USA).

2.10.2 High Resolution Melt (HRM) marker

HRM markers for genes of interest were developed based on sequence differences between alleles. Primers were designed to amplify small fragments (<200-bp) that contained these allelic polymorphisms. HRM markers were scored using the Rotor-Gene Q machine (Corbett Research, Australia). The CAS-1200™ pipetting robot (Corbett Research, Australia) with CAS Robotics v4.9.8 software was employed to prepare reactions containing 2µL template, 1.05µL forward primer, 1.05µL reverse primer, 7.5µL HRM PCR Master Mix from Qiagen HRM PCR Kit, and 3.4µL autoclaved Milli-Q water. HRM reactions were performed for 50 cycles using the following temperature configurations: 95°C for 5 min, single cycle (95°C for 10 seconds, 58°C for 30 seconds), 95°C for 5 min, 50°C for 5 min, HRM (0.1°C temperature increments from 60-90°C, or from product melt temperature -5°C to +5°C). HRM results were analysed using the Qiagen ScreenClust HRM® Software.

2.10.3 Kompetitive Allele Specific PCR (KASP) marker

A primer mix containing two forward allele-specific primers and a single reverse common primer, were designed for allelic polymorphisms between parental accessions. Each forward primer is additionally designed with a tail sequence that binds to one of two fluorescent resonance energy transfer cassettes included in the KASP assay mix.

KASP assays were carried out at the University of Saskatchewan. Refer to manufacturer's instructions (http://www.lgcgroup.com/products/kasp-genotyping-chemistry/) for more information.

2.10.4 Allele-specific PCR marker

The allele-specific PCR markers were designed to genotype parental accessions that had large deletions (>1000-bp) for genes and intergenic regions of interest. The PCR mix comprised of three primers, two allele-specific primers and a single common primer. Primers were designed to amplify DNA fragments of varying sizes, in order to obtain visible bands that were allele-specific for each parental accession. PCR reactions were conducted as per Standard PCR (section 2.6.1).

2.11 Linkage and Quantitative Trait Loci (QTL) analysis

Refer to individual chapters for more information.

2.12 Construction of sequence alignments

Amino acid sequences of predicted proteins were aligned using ClustalX (Thompson et al., 1997) and adjusted manually where appropriate using GeneDoc v2.7.000 (Nicholas and Nicholas, 1997). Percentage identity for homologous predicted proteins were calculated in GeneDoc from full-length protein alignments constructed using ClustalX.

2.13 Statistical analysis

All statistical analyses were conducted using StatPlus®:mac LE by AnalystSoft Inc. A significance level of 0.05 was employed for all analyses.

Chapter 3 The molecular basis for the Lentil *Sn* locus

3.1 Introduction

The genetic characterisation of lentil *Sn* by Sarker et al. (1999) is the first significant study on the genetic control of flowering time in lentil. The lentil *Sn* was first described in cv. Precoz (Sarker et al., 1999), an early-flowering accession that has been utilised extensively in breeding programs to widen the genetic pool of the *pilosae* germplasm (Erskine et al., 1998; Rahman et al., 2009). This locus was designated *Sn* by analogy with the locus of the same name in pea (Murfet, 1971), although no evidence supporting the orthology of these loci has been presented (Sarker et al., 1999).

3.1.1 Origins of lentil Sn

The early-flowering cv. Precoz was first reported to occur in Argentina, and is documented to have been cultivated in the major lentil production regions of the country since 1967-68 (Riva, 1975). The genetic affinity of this accession to the global germplasm however remains undetermined. From the early work by Barulina (1930) it can be inferred that cv. Precoz is likely derived from the *macrosperma* form first introduced by the Spanish into South America. Barulina (1930) had also then described these introduced lentils to have a narrow genetic base. When evaluated in Argentina, Riva (1975) reported that no other accessions demonstrated comparable precocity to cv. Precoz. The precocity of cv. Precoz was desirable to growers as it prevented the large outbreaks of *Fusarium oxysporum* sp*. lentis*, *Fusarium roseum* var. gibbosum and *Rhizoctonia solani* Kuhn (Riva, 1975).

3.1.2 Characterisation of lentil Sn locus

The current literature presents a limited understanding of the lentil *Sn*. Beyond the initial characterisation of its role in conferring a recessive early-flowering

phenotype, and the reported linkage of *Sn* to loci controlling seed coat protein (*Scp*) and peduncle pubescence (*Pep*) as reported by Sarker et al. (1999), little is known about the locus.

While the characterisation of the lentil *Sn* has been limited, studies on cv. Precoz have allowed a broader appreciation of the flowering time locus. Studies on flowering time in lentil by Roberts et al. (1986) and Summerfield et al (1985) determined cv. Precoz to be vernalisation responsive, and to demonstrate reduced sensitivity to photoperiod. The early-flowering accession is also described to have an erect (*ert*), non-bushy habit with moderate branching (Emami and Sharma, 1999). *Ert* is linked with loci that regulate anthocyanin pigmentation in stems (*Gs*) and leaves (*Bl*) (Emami and Sharma, 1999). Separately, both *Ert* and *Scp* are defined to be linked to a locus controlling pod dehiscence (*Pi*) (Kumar et al., 2005; Tahir and Muehlbauer, 1994).

3.1.3 Significance of lentil Sn in current breeding programs

The lentil *Sn* has been crucial to breaking the reproductive isolation of the *pilosae* lentil imposed by the asynchrony in flowering between West Asian accessions and the *pilosae* germplasm (Erskine et al., 1998). Early alleles for the lentil *Sn* facilitated the introduction of new genetic material from exotic accessions like cv. Precoz into breeding programs that aimed to broaden the narrow genetic base of the *pilosae* lentil. Erskine et al. (1998) has suggested that cv. Precoz and its derivatives are included in every crossing block of the Indian breeding program.

3.1.4 Chapter aims

This chapter aims to determine the physiological and molecular basis for the early-flowering habit conferred by lentil *Sn*. The chapter also seeks to understand the pleiotropic effect of the proposed candidate on key agronomic traits, including branching, internode length, and germination time. The chapter additionally seeks to determine the prevalence of the recessive early-flowering allele in a representative collection of cultivated lentil.

3.2 Materials and methods

This section details specific materials and methods relevant to this chapter. General materials and methods are described in Chapter 2.

3.2.1 Plant materials and growth conditions

A F_2 population derived from a single cross between ILL 6005 and ILL 5588 (cv. Northfield) was evaluated in this chapter. ILL 6005 is an early flowering selection derived from a cross between cv. Precoz and cv. Laird (Erskine, W. pers. comm.), and is reported to carry the recessive form of the lentil *Sn* (Weller and Murfet, unpublished).

A F_2 population derived from a single cross between ILL 223 and ILL 5588 was also evaluated in this chapter. ILL 223 is an Iranian accession that carries a 3-bp indel in the coding sequence of the lentil *ELF3* gene. ILL 223 was obtained from the Australian Temperate Field Crops Collection (ATFCC).

A University of Saskatchewan Lentil Association Mapping (LAM) panel, comprised of 94 lentil accessions, was analysed in this chapter. Details of the accessions are summarised in Appendix 1.

Plants were evaluated at the University of Tasmania phytotron. Details of photoperiod conditions for each experiment in this chapter are summarised in Table 3-1.

3.2.2 Molecular markers & genotyping

Single nucleotide polymorphism (SNP), High Resolution Melt (HRM), and Kompetitive Allele Specific PCR (KASP) markers were utilised in this chapter. Molecular markers were developed for both *elf3-1* and *elf3-2* allelic variants of *LcELF3* for both ILL 6005 x ILL 5588 and ILL 223 x ILL 5588 F_2 populations. Molecular markers for *LcMYB1, LcELF4, LcLUX, LcTOC1, LcPRR37,* and *LcPRR59a* were also developed for co-segregation analysis in the ILL 6005 x ILL 5588 F_2 population. Molecular marker information are summarised in Table 3-1.

Experiment Description	Treatment	Conditions	Figure
Phenotypic characterisation of *Sn* under different photoperiods	LD	8-h natural daylight, extended with 8-h low irradiance (10 µmol·m^{-2}s^{-1}) white light from mixed florescent & incandescent sources.	Figure 3-1
	SD	8-h natural daylight, extended with 2-h low irradiance (10 µmol·m^{-2}s^{-1}) white light from mixed florescent & incandescent sources.	
Phenotypic characterisation of *Sn* under continuous monochromatic light	Red	Continuous red light (20 µmol m^{-2} s^{-1}); 3 weeks.	Figure 3-2
	Far-red	Continuous white light (20 µmol m^{-2} s^{-1}); 3 weeks.	
	Blue	Continuous blue light (20 µmol m^{-2} s^{-1}); 3 weeks.	
	Dark	Continuous darkness; 3 weeks.	
	White	Continuous white light (110 µmol m^{-2} s^{-1}); 3 weeks.	
Segregation analysis of ILL 6005 x ILL 5588 F_2 population	SD	8-h natural daylight, extended with 2-h low irradiance (10 µmol·m^{-2}s^{-1}) white light from mixed florescent & incandescent sources.	Figure 3-3, Figure 3-4
Phenotypic characterisation of ILL 223 (*elf3-2*) under different photoperiods	LD	12-h natural daylight, extended with 4-h low irradiance (10 µmol·m^{-2}s^{-1}) white light from mixed florescent & incandescent sources.	Figure 3-8
	SD	12-h natural daylight.	
Segregation analysis of ILL 223 x ILL 5588 F_2 population	SD	12-h natural daylight.	Figure 3-8

Table 3-1 Summary of photoperiod and light conditions for experiments.

3.2.3 Plant measurements

Refer to Chapter 2.

3.3 Results

3.3.1 Phenotypic characterisation of Sn under different photoperiods

The physiological basis for the early-flowering phenotype conferred by the lentil *Sn* is not known. A phenotypic evaluation was carried out under controlled long day (LD) and short day (SD) photoperiods to characterise the photoperiod response of ILL 6005. Observations for this evaluation were made relative to the photoperiod-sensitive, medium-late flowering accession ILL 5588.

ILL 6005 was evaluated to flower earlier, scored as *days to flowering* (DTF) and *node of flower development* (NFD), than ILL 5588 in both LD and SD (Figure 3-1). There was no effect of photoperiod on DTF for ILL 6005, suggesting that the cv. Precoz derivative is photoperiod-insensitive (Figure 3-1A). In contrast, ILL 5588 was observed to flower later under SD when compared plants to LD, affirming its photoperiod-sensitivity (Figure 3-1).

ILL 5588 was also evaluated to demonstrate floral abortions prior to the development of the first open flower under SD. No floral abortions were reported in ILL 6005 or in LD exposed ILL 5588. Floral abortions observed in ILL 5588 were not formally scored in this chapter. This phenomenon is further explored in Chapter 4 (refer to section 4.3.1).

The flowering time observations are summarised in Table 3-2.

		LD		SD		*p*-value	Figure
		mean	± SE	mean	± SE		
ILL 6005	**DTF** (days)	32.4	1.30	34.7	0.667	0.1525	Figure 3-1A
	NFD (nodes)	12.0	0.258	13.8	0.147	0.0000*	Figure 3-1B
ILL 5588	**DTF** (days)	41.3	1.41	81.0	0.745	0.0000*	Figure 3-1A
	NFD (nodes)	17.0	0.258	26.2	0.629	0.0000*	Figure 3-1B

Table 3-2 Summary of DTF and NFD under different photoperiods
Days to flowering (DTF) and Node of flower development (NFD) in ILL 6005 and ILL 5588 under LD and SD conditions. Plants received an 8-h photoperiod of natural daylight, extended with 2-h low irradiance (SD) and 8-h low irradiance (LD) of white light from mixed florescent & incandescent sources. Asterisk (*) indicates statistical significance ($p < 0.05$). Data are mean ±SE for n=9-10.

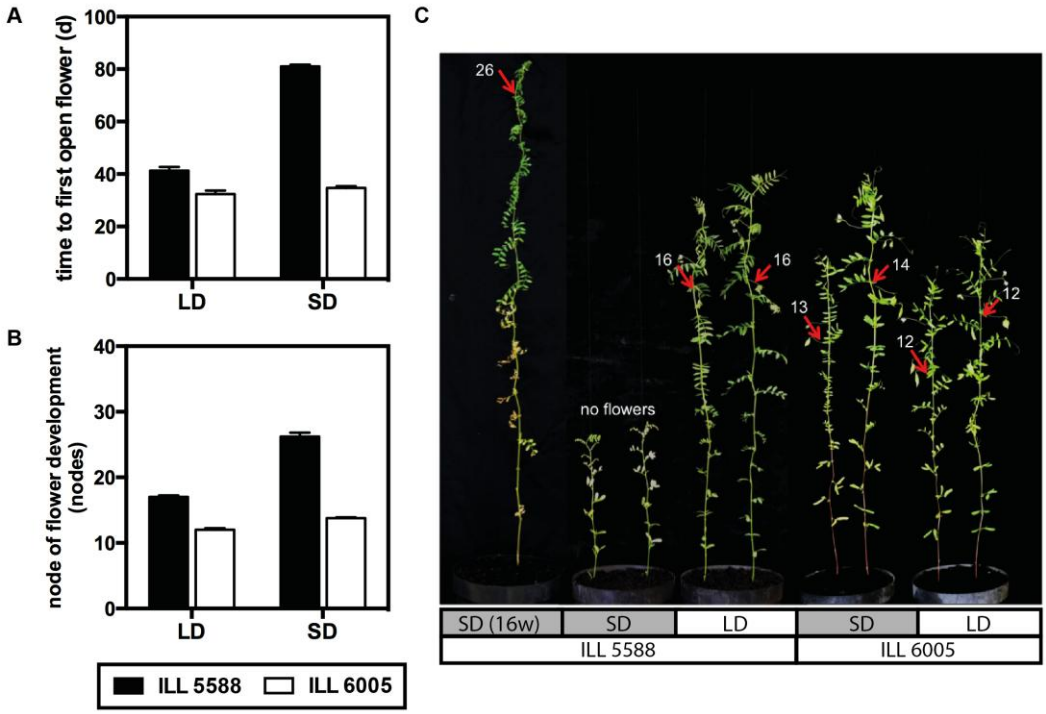


Figure 3-1 Phenotypic characterisation of *Sn* under different photoperiods.
(A) Days to flowering (DTF) in ILL 5588 (*Sn*) and ILL 6005 (*sn*), under LD[1] and SD[1] conditions. (B) Node of flower development (NFD) in ILL 5588 (*Sn*) and ILL 6005 (*sn*), under LD and SD conditions. (C) Representatives of ILL 5588 (*Sn*) and ILL 6005 (*sn*) (with lateral branches excised) grown under LD[2] and SD[2] conditions at 7-weeks from emergence, and ILL 5588 (with lateral branches excised) grown under SD at 16-weeks from emergence. Red arrows and numbers denote NFD. [1]Plants received an 8-h photoperiod of natural daylight, extended with 4-h (SD) or 8-h (LD) low irradiance (10 $\mu mol \cdot m^{-2} \cdot s^{-1}$) white light from mixed florescent and incandescent sources. [2]Plants received a 12-h photoperiod of natural daylight (SD) and a 12-h photoperiod of natural daylight extended with 4-hours of fluorescent light (LD). Data are mean ±SE for *n*=9-10.

3.3.2 Evaluating the role of photoreceptors in conferring photoperiod-insensitivity to ILL 6005

In the phylogenetically related *Pisum sativum* (Weller et al., 2004), and in *Arabidopsis* (Mockler et al., 1999), defects in light perception and light signal transduction have been associated with a photoperiod-insensitive early flowering phenology. Defects in photoreceptors also manifests in abnormal seedling photomorphogenesis, including elongated internodes.

To investigate if photoreceptors have a role in conferring ILL 6005 photoperiod-insensitivity, a preliminary experiment evaluating the internode length of seedlings exposed to continuous monochromatic light from sowing was carried out.

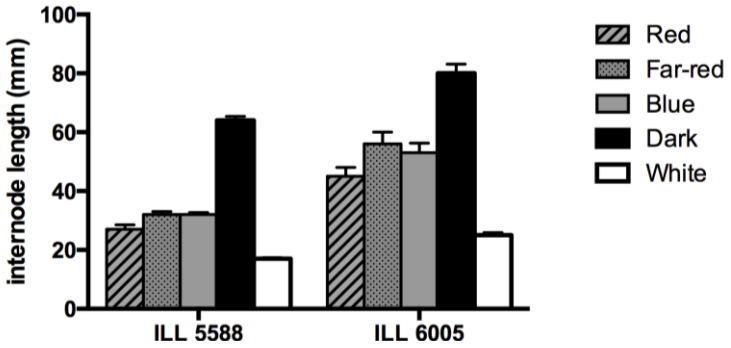


Figure 3-2 Early-flowering ILL 6005 under continuous monochromatic light.
Stem elongation in ILL 5588 and ILL 6005 seedling under continuous red, far-red, blue monochromatic light and in white light and in darkness for 2-weeks from sowing. Internode length was measured as length between nodes 1 and 3. Data are mean ±SE for *n*=3-9.

It was observed that ILL 6005 displayed longer internodes under all conditions, when compared to ILL 5588. However there was no monochromatic condition that resulted in discernably abnormal elongation (Figure 3-2). While the differences in the genetic backgrounds of ILL 6005 and ILL 5588 limited direct comparisons or further analyses, it can be inferred from this preliminary study that defects in photoreceptors are unlikely to contribute to the observed photoperiod-insensitivity of ILL 6005.

3.3.3 Segregation of ILL 6005 x ILL 5588 F_2 population for flowering time

To further evaluate the basis for the photoperiod-insensitive early-flowering phenology conferred by the lentil *Sn*, a cross between ILL 6005 and ILL 5588 was generated and F_2 progeny evaluated under a controlled 10-h photoperiod. While a 10-h photoperiod is shorter than the shortest days experienced at the lower end of the latitudinal range for lentil cultivation, it was predicted that a shorter photoperiod would accentuate the difference in flowering time between the lentil *Sn* and *sn* genotypes.

Bimodality in the segregation for flowering time with an early class and late class was observed (Figure 3-3A). The early segregants flowered on average in 34.5 ± 1.52 days and at node 13.3 ± 0.336, while the late segregants flowered on average in 78.4 ± 1.30 days and at node 21.7 ± 0.540. A moderate positive correlation (R^2_{adj} = 0.553) was determined between DTF and NFD.

The 3:1 Mendelian nature (p = 0.178) of the segregation points to the dominance of the late flowering phenotype, and implies that a single locus controlling flowering time is probable. These observations are consistent with findings by Sarker et al. (1999), suggesting that it is likely the F_2 population is segregating for lentil *Sn*.

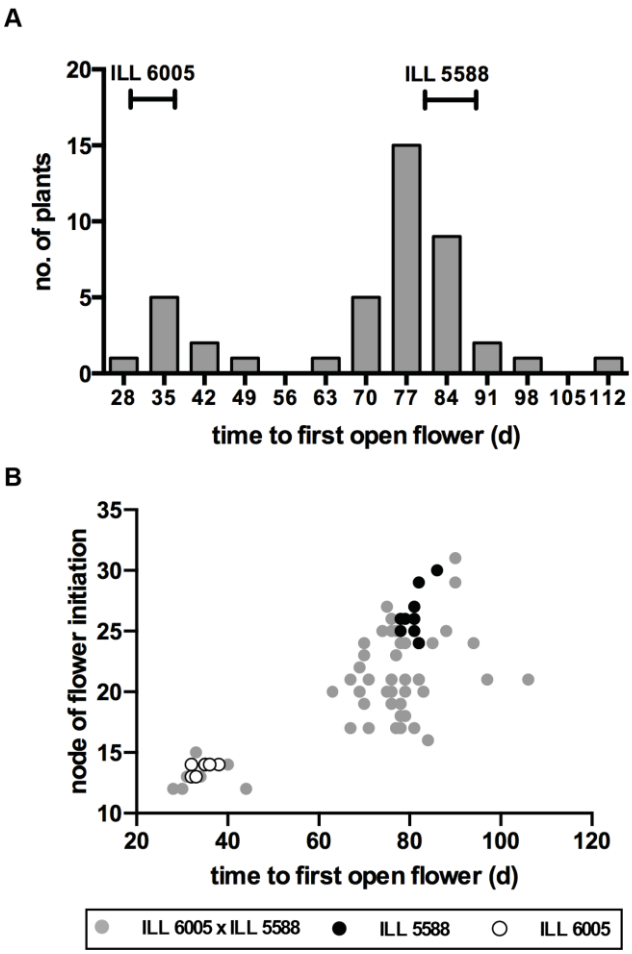


Figure 3-3 Segregation of ILL 6005 x ILL 5588 F_2 population for flowering time.
(A) Bimodal distribution F_2 progeny with respect to flowering time under SD conditions. (B) Transition to reproductive development in F_2 progeny illustrated in the context of days to flowering (DTF) and node of flower development (NFD) under SD conditions. Data for (A) and (B) are n=9-53. Plants received an 8-h photoperiod of natural daylight, extended with 2-h low irradiance (10 $\mu mol \cdot m^{-2} \cdot s^{-1}$) white light from mixed florescent and incandescent sources.

3.3.4 Genetic evaluation of candidate genes for Sn

A candidate-gene approach then was employed to determine the genetic basis for the observed bimodal segregation for flowering time in the F_2 population. Previous work on the genetic control of flowering time in *P. sativum*, as described in Section 1.6, allows for a comparative approach to candidate-gene determination.

In *P. sativum*, mutants for *HR* (Weller et al., 2012), *SN* (Liew et al., 2014) and *DNE* (Liew et al., 2009a), all components of the circadian clock, have also been reported to display photoperiod-insensitivity and are described to be early-flowering. To facilitate a candidate-gene analysis for the F_2 population, partial sequences of lentil orthologues for *M. truncatula* and *P. sativum* circadian clock components were isolated, and SNP and HRM markers designed to distinguish the ILL 6005 and ILL 5588 alleles for the circadian clock components (Table 3-3). Refer to Appendix 2 for marker information and Appendix 3 for primer information.

Gene symbol	Medicago locus (Mt4.0)	Gene description	Additional information
ELF3	Medtr3g103970	circadian clock component	*Arabidopsis ELF3* mutants are early-flowering (Zagotta et al., 1996). Early-flowering mutants described in *P. sativum* (Weller et al., 2012), *H. vulgare* (Faure et al., 2012), and *O. sativa* (Matsubara et al., 2012). *ELF3* is a circadian clock component and interacts with other flowering time genes (Kolmos et al., 2011; Liu et al., 2001; Lu et al., 2012; Nefissi et al., 2011; Nusinow et al., 2011; Serrano, 2011; Yu et al., 2008).
LHY/ MYB1	Medtr7g146190	late elongated hypocotyl-like protein	*P. sativum MYB1* demonstrates strong diurnal expression rhythm in LD and SD. *P. sativum MYB1* is a *Arabidopsis CCA1/LHY* orthologue (Hecht et al., 2007; Liew et al., 2009a).
LUX	Medtr4g064730	circadian clock component	*Arabidopsis LUX* is an evening component of the circadian clock (Helfer et al., 2011; Nusinow et al., 2011). *P. sativum SN* is a *LUX* orthologue and recessive alleles confer early-flowering habit (Liew et al., 2014).
PRR37	Medtr4g079920	Pseudo-Response Regulator	*P. sativum PRR37 demonstrates strong du* diurnal expression rhythm in LD and SD (Liew et al., 2009a). *Arabidopsis PRR* mutants demonstrate defects in photoperiod control of flowering time (Nakamichi et al., 2007).
PRR59a	Medtr3g092780	Pseudo-Response Regulator	*P. sativum PRR59a* demonstrates strong diurnal expression rhythm in LD and SD (Liew et al., 2009a). *Arabidopsis PRR* mutants demonstrate defects in photoperiod control of flowering time (Nakamichi et al., 2007).
TOC1	unknown	Pseudo-Response Regulator	*Arabidopsis TOC1* mutants demonstrate early flowering (Sato et al., 2002). *P. sativum TOC1 demonstrates strong du* diurnal expression rhythm in LD and SD (Liew et al., 2009a)

Table 3-3 Summary of candidate genes

One of these identified candidates *LcELF3*, an orthologue of the *Arabidopsis EARLY FLOWERING 3* circadian clock gene (Zagotta et al., 1996), showed perfect co-segregation with the lentil *Sn* phenotype, with all early-flowering segregants being homozygous for the ILL 6005 allele, and all late segregants carrying at least one ILL 5588 allele (Figure 3-4).

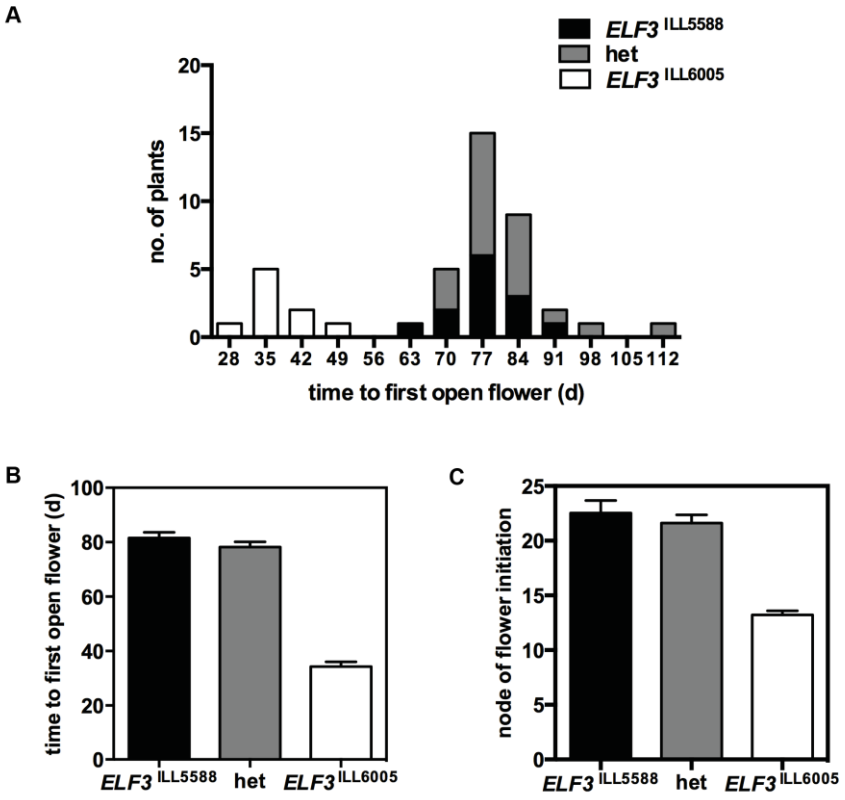


Figure 3-4 Genetic association of *LcELF3* to flowering time in ILL 6005 x ILL 5588 F_2 population. (A) Co-segregation of early-flowering phenotype with *LcELF3* under SD. (B) Association analysis of *LcELF3* with flowering time under SD. Data are mean ±SE for *n*=9-21. (C) Association analysis of *LcELF3* with flowering node under SD conditions. Data are mean ±SE for *n*=9-21. Plants received an 8-h photoperiod of natural daylight, extended with 2-h low irradiance (10 $\mu mol \cdot m^{-2} \cdot s^{-1}$) white light from mixed florescent and incandescent sources.

Segregants homozygous for the ILL 6005 allele flowered on average in 34.3 ± 1.69 days and at node 13.2 ± 0.364, significantly earlier when compared to segregants that were either heterozygous (DTF = 78.2 ± 1.95, NFD = 21.6 ± 0.742; $p < 0.05$), or homozygous for the ILL 5588 allele (DTF = 81.5 ± 2.05, NFD = 22.5 ± 1.14; $p < 0.05$) (Figure 3-4B and Figure 3-4C). There was no significant difference for DTF or NFD between segregants that were heterozygous (p = 0.266) or homozygous (p = 0.484) for the ILL 5588 allele.

3.3.5 Molecular evaluation of LcELF3 as a candidate for Sn

The complete co-segregation of the early-flowering phenotype with *LcELF3* suggests that *LcELF3* is tightly linked to the QTL responsible for the observed phenotype, and that the late-flowering allele is dominant. The latter is consistent with the existing perspective on the lentil *Sn*; that the early-flowering habit is conferred only in its recessive state.

To evaluate the possibility that the lentil *Sn* locus is equivalent to *LcELF3*, the full-length of the genomic and coding sequence for the lentil orthologue was isolated and sequenced in ILL6005 and ILL5588, and in the parents of ILL6005 (cv. Laird and cv. Precoz). The genomic and coding sequences from each accession were then analysed for polymorphisms that had the potential to contribute to the observed early-flowering phenotype.

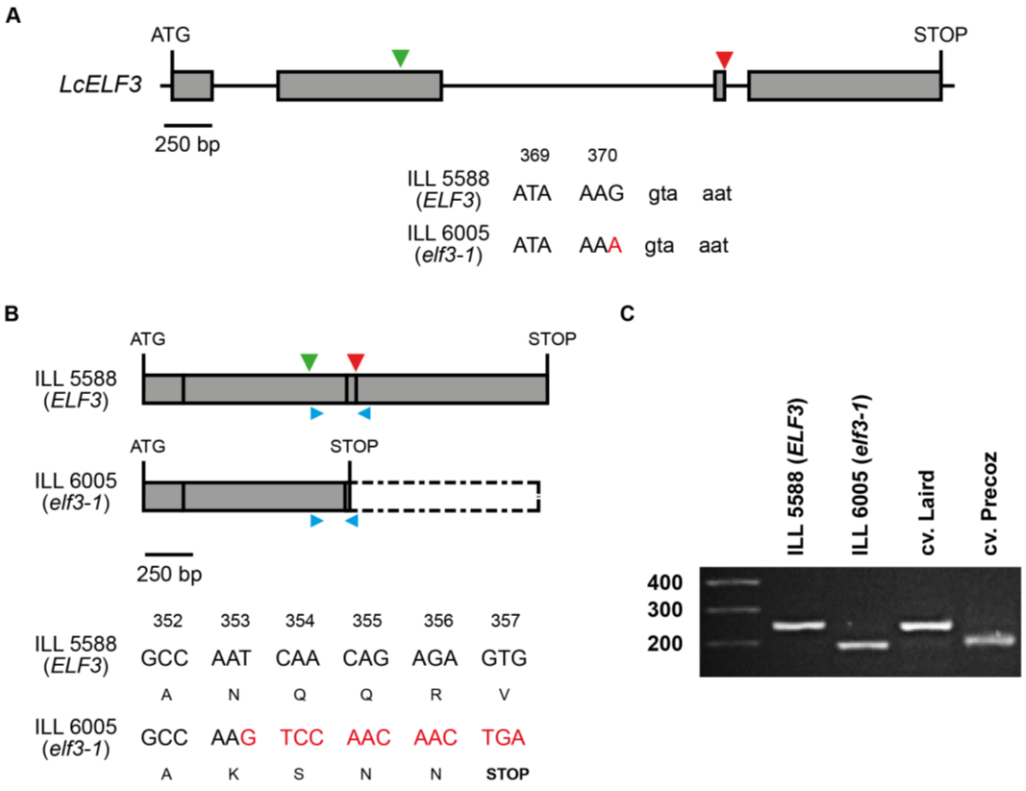


Figure 3-5 Nature of polymorphisms in ILL 6005.

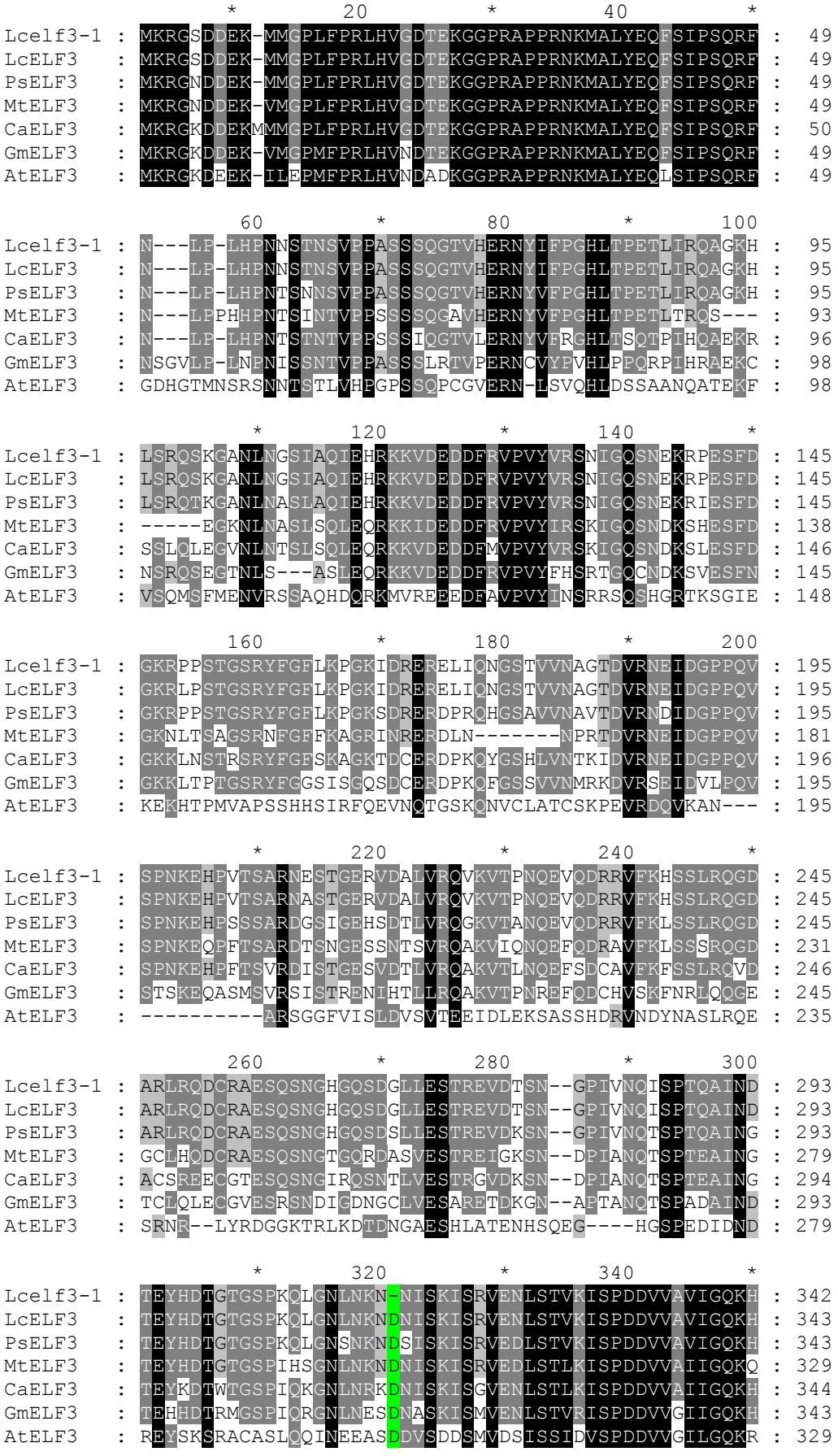
(A) Position of G-to-A substitution (red) and 3-bp indel (green), and details of G-to-A substitution in the *LcELF3* genomic sequence. (B) Position of G-to-A substitution (red) and 3-bp indel (green), and details of the splicing defect in the *LcELF3* mRNA and predicted truncation of LcELF3 protein.. (C) PCR revealing the 52-bp deletion in *LcELF3* mRNA, corresponding to the skipping of exon 3. Blue arrows indicate position of PCR primers. Refer to Appendix 3 for primer details.

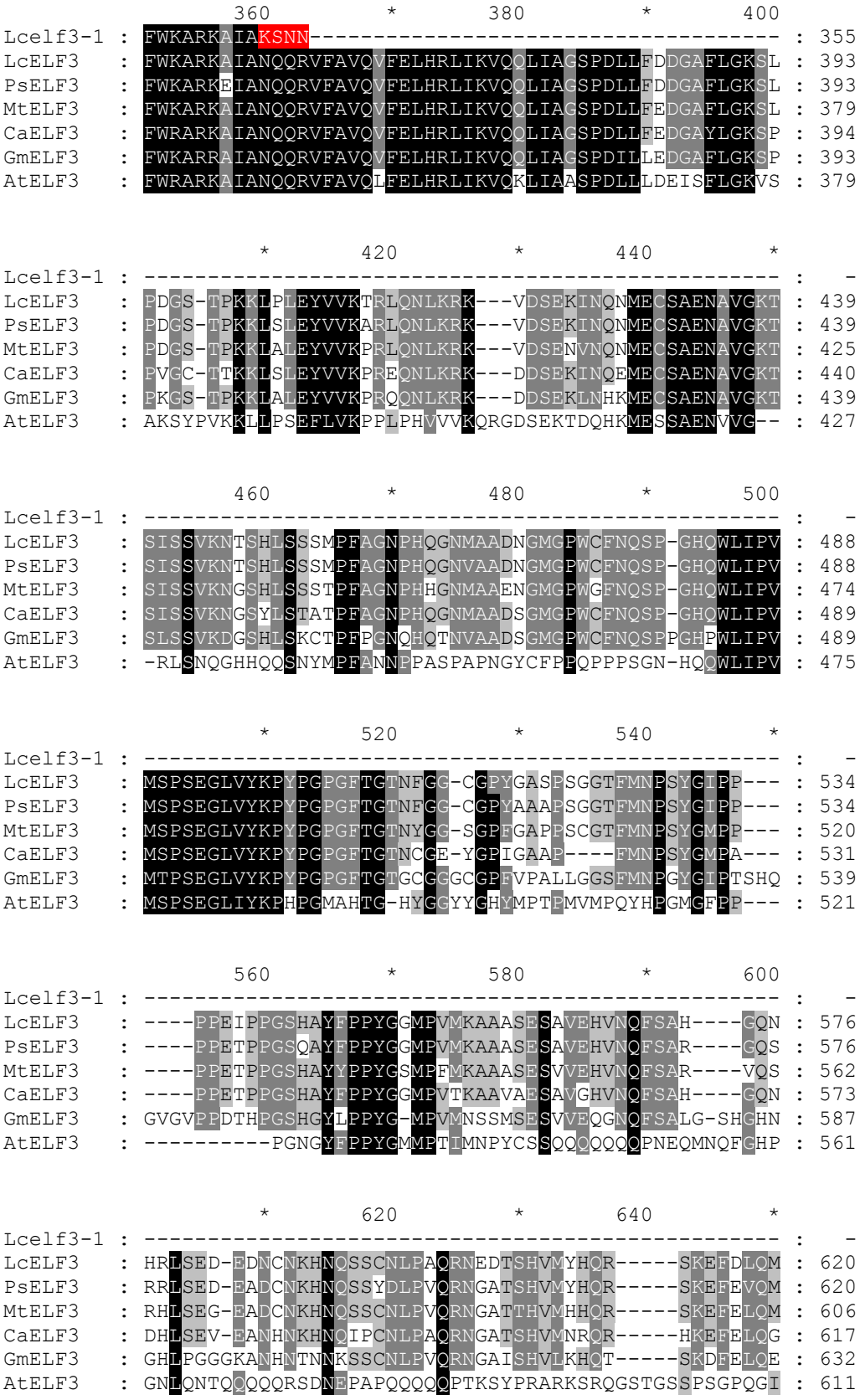
It was observed that genomic *LcELF3* from both ILL 6005 and cv. Precoz had a translationally silent G-to-A substitution in the last nucleotide of exon 3, at the exon 3-intron 3 recognition site, as indicated by the red arrow in Figure 3-5A. This substitution was observed to result in the missplicing and subsequent skipping of exon 3 of the coding sequence. This missplicing is predicted to cause a frame-shift, and a premature stop-codon after four missense amino acids during translation in the coding sequence (Figure 3-6). The missplicing and exon skipping was further verified by PCR amplification of the region between exons 2 and 4, as indicated by blue arrows in Figure 3-5B, which revealed a 52-bp deletion in the mRNA of *LcELF3* in both ILL 6005 and cv. Precoz (Figure 3-5C).

Additionally, a 3-bp deletion in exon 2 that results in the predicted deletion of an aspartic acid, as illustrated by the green arrow in Figure 3-5A was also noted. The deleted aspartic acid is conserved across *P. sativum*, *Medicago truncatula*, *Cicer arietinum, Glycine max*, and *Arabidopsis* (Figure 3-6). It is unclear if this polymorphism has any functional significance. No other polymorphisms with predicted amino acid changes were observed.

The tight linkage observed between *ELF3* and the lentil *Sn* early-flowering, and the predicted truncation of the ELF3 protein from the early flowering ILL 6005 and cv. Precoz suggests that *LcELF3* allele derived from cv. Precoz is likely responsible for the observed early-flowering phenotype. This is supported by the derived role of the circadian clock gene and the photoperiod-insensitive early-flowering phenotype reported by *ELF3* mutants in *Arabidopsis* (Lu et al., 2012; Zagotta et al., 1996), phylogenetically related *P. sativum* (Weller et al., 2012), and barley (Faure et al., 2012).

This chapter proposes that the early-flowering allele be designated *elf3-1*.


Lcelf3-1 : MKRGSDDEK-MMGPLFPRLHVGDTEKGGPRAPPRNKMALYEQFSIPSQRF : 49
LcELF3 : MKRGSDDEK-MMGPLFPRLHVGDTEKGGPRAPPRNKMALYEQFSIPSQRF : 49
PsELF3 : MKRGNDDEK-MMGPLFPRLHVGDTEKGGPRAPPRNKMALYEQFSIPSQRF : 49
MtELF3 : MKRGNDDEK-VMGPLFPRLHVGDTEKGGPRAPPRNKMALYEQFSIPSQRF : 49
CaELF3 : MKRGKDDEKMMMGPLFPRLHVGDTEKGGPRAPPRNKMALYEQFSIPSQRF : 50
GmELF3 : MKRGKDDEK-VMGPMFPRLHVNDTEKGGPRAPPRNKMALYEQFSIPSQRF : 49
AtELF3 : MKRGKDEEK-ILEPMFPRLHVNDADKGGPRAPPRNKMALYEQLSIPSQRF : 49
Lcelf3-1 : N---LP-LHPNNSTNSVPPASSSQGTVHERNYIFPGHLTPETLIRQAGKH : 95
LcELF3 : N---LP-LHPNNSTNSVPPASSSQGTVHERNYIFPGHLTPETLIRQAGKH : 95
PsELF3 : N---LP-LHPNTSNNSVPPASSSQGTVHERNYVFPGHLTPETLIRQAGKH : 95
MtELF3 : N---LPPHHPNTSINTVPPSSSSQGAVHERNYVFPGHLTPETLTRQS--- : 93
CaELF3 : N---LP-LHPNTSTNTVPPSSSIQGTVLERNYVFRGHLISQTPIHQAEKR : 96
GmELF3 : NSGVLP-LNPNISSNTVPPASSSLRTVPERNCVYPVHLPPQRPIHRAEKC : 98
AtELF3 : GDHGTMNSRSNNTSTLVHPGPSSQPCGVERN-LSVQHLDSSAANQATEKF : 98
Lcelf3-1 : LSRQSKGANLNGSIAQIEHRKKVDEDDFRVPVYVRSNIGQSNEKRPESFD : 145
LcELF3 : LSRQSKGANLNGSIAQIEHRKKVDEDDFRVPVYVRSNIGQSNEKRPESFD : 145
PsELF3 : LSRQTKGANLNASLAQIEHRKKVDEDDFRVPVYVRSNIGQSNEKRIESFD : 145
MtELF3 : -----EGKNLNASISQLEQRKKIDEDDFRVPVYIRSKIGQSNDKSHESFD : 138
CaELF3 : SSLQLEGVNLNTSISQLEQRKKVDEDDFMVPVYVRSKIGQSNDKSLESFD : 146
GmELF3 : NSRQSEGTNLS---ASLEQRKKVDEDDFRVPVYFHSRTGQCNDKSVESFN : 145
AtELF3 : VSQMSFMENVRSSAQHDQRKMVREEEDFAVPVYINSRRSQSHGRTKSGIE : 148
Lcelf3-1 : GKRPPSTGSRYFGFLKPGKIDRERELIQNGSTVVNAGTDVRNEIDGPPQV : 195
LcELF3 : GKRLPSTGSRYFGFLKPGKIDRERELIQNGSTVVNAGTDVRNEIDGPPQV : 195
PsELF3 : GKRPPSTGSRYFGFLKPGKSDRERDPRQHGSAVVNAVTDVRNDIDGPPQV : 195
MtELF3 : GKNLTSAGSRNFGFFKAGRINRERDLN-------NPRTDVRNEIDGPPQV : 181
CaELF3 : GKKLNSTRSRYFGFSKAGKTDCERDPKQYGSHLVNTKIDVRNEIDGPPQV : 196
GmELF3 : GKKLTPTGSRYFGGSISGQSDCERDPKQFGSSVVNMRKDVRSEIDVLPQV : 195
AtELF3 : KEKHTPMVAPSSHHSIRFQEVNQTGSKQNVCLATCSKPEVRDQVKAN--- : 195
Lcelf3-1 : SPNKEHPVTSARNESTGERVDALVRQVKVTPNQEVQDRRVFKHSSLRQGD : 245
LcELF3 : SPNKEHPVTSARNASTGERVDALVRQVKVTPNQEVQDRRVFKHSSLRQGD : 245
PsELF3 : SPNKEHPSSSARDGSIGEHSDTLVRQGKVTANQEVQDRRVFKLSSLRQGD : 245
MtELF3 : SPNKEQPFTSARDTSNGESSNTSVRQAKVIQNQEFQDRAVFKLSSSRQGD : 231
CaELF3 : SPNKEHPFTSVRDISTGESVDTLVRQAKVTLNQEFSDCAVFKFSSLRQVD : 246
GmELF3 : STSKEQASMSVRSISTRENIHTLLRQAKVTPNREFQDCHVSKFNRLQQGE : 245
AtELF3 : ----------ARSGGFVISLDVSVTEEIDLEKSASSHDRVNDYNASLRQE : 235
Lcelf3-1 : ARLRQDCRAESQSNGHGQSDGLLESTREVDTSN--GPIVNQISPTQAIND : 293
LcELF3 : ARLRQDCRAESQSNGHGQSDGLLESTREVDTSN--GPIVNQISPTQAIND : 293
PsELF3 : ARLRQDCRAESQSNGHGQSDSLLESTREVDKSN--GPIVNQTSPTQAING : 293
MtELF3 : GCLHQDCRAESQSNGTGQRDASVESTREIGKSN--DPIANQTSPTEAING : 279
CaELF3 : ACSREECGTESQSNGIRQSNTLVESTRGVDKSN--DPIANQTSPTEAING : 294
GmELF3 : TCLQLECGVESRSNDIGDNGCLVESARETDKGN--APTANQTSPADAIND : 293
AtELF3 : SRNR--LYRDGGKTRLKDTDNGAESHLATENHSQEG----HGSPEDIDND : 279
Lcelf3-1 : TEYHDTGTGSPKQLGNLNKN-NISKISRVENLSTVKISPDDVVAVIGQKH : 342
LcELF3 : TEYHDTGTGSPKQLGNLNKNDNISKISRVENLSTVKISPDDVVAVIGQKH : 343
PsELF3 : TEYHDTGTGSPKQLGNSNKNDSISKISRVEDLSTVKISPDDVVAVIGQKH : 343
MtELF3 : TEYHDTGTGSPIHSGNLNKNDNISKISRVEDLSTLKISPDDVVAIIGQKQ : 329
CaELF3 : TEYKDTWTGSPIQKGNLNRKDNISKISGVENLSTLKISPDDVVAIIGQKH : 344
GmELF3 : TEHHDTRMGSPIQRGNLNESDNASKISMVENLSTVRISPDDVVGIIGQKH : 343
AtELF3 : REYSKSRACASLQQINEEASDDVSDDSMVDSISSIDVSPDDVVGILGQKR : 329


360 * 380 * 400
Lcelf3-1 : FWKARKAIAKSNN---------- : 355
LcELF3 : FWKARKAIANQQRVFAVQVFELHRLIKVQQLIAGSPDLLFDDGAFLGKSL : 393
PsELF3 : FWKARKEIANQQRVFAVQVFELHRLIKVQQLIAGSPDLLFDDGAFLGKSL : 393
MtELF3 : FWKARKAIANQQRVFAVQVFELHRLIKVQQLIAGSPDLLFEDGAFLGKSL : 379
CaELF3 : FWRARKAIANQQRVFAVQVFELHRLIKVQQLIAGSPDLLFEDGAYLGKSP : 394
GmELF3 : FWKARRAIANQQRVFAVQVFELHRLIKVQQLIAGSPDILLEDGAFLGKSP : 393
AtELF3 : FWRARKAIANQQRVFAVQLFELHRLIKVQKLIAASPDLLLDEISFLGKVS : 379
* 420 * 440 *
Lcelf3-1 : ---------- : -
LcELF3 : PDGS-TPKKLPLEYVVKTRLQNLKRK---VDSEKINQNMECSAENAVGKT : 439
PsELF3 : PDGS-TPKKLSLEYVVKARLQNLKRK---VDSEKINQNMECSAENAVGKT : 439
MtELF3 : PDGS-TPKKLALEYVVKPRLQNLKRK---VDSENVNQNMECSAENAVGKT : 425
CaELF3 : PVGC-TTKKLSLEYVVKPREQNLKRK---DDSEKINQEMECSAENAVGKT : 440
GmELF3 : PKGS-TPKKLALEYVVKPRQQNLKRK---DDSEKLNHKMECSAENAVGKT : 439
AtELF3 : AKSYPVKKLLPSEFLVKPPLPHVVVKQRGDSEKTDQHKMESSAENVVG-- : 427
460 * 480 * 500
Lcelf3-1 : ---------- : -
LcELF3 : SISSVKNTSHLSSSMPFAGNPHQGNMAADNGMGPWCFNQSP-GHQWLIPV : 488
PsELF3 : SISSVKNTSHLSSSMPFAGNPHQGNVAADNGMGPWCFNQSP-GHQWLIPV : 488
MtELF3 : SISSVKNGSHLSSSTPFAGNPHHGNMAAENGMGPWGFNQSP-GHQWLIPV : 474
CaELF3 : SISSVKNGSYLSTATPFAGNPHQGNMAADSGMGPWCFNQSP-GHQWLIPV : 489
GmELF3 : SLSSVKDGSHLSKCTPFPGNQHQTNVAADSGMGPWCFNQSPPGHPWLIPV : 489
AtELF3 : -RLSNQGHHQQSNYMPFANNPPASPAPNGYCFPPQPPPSGN-HQQWLIPV : 475
* 520 * 540 *
Lcelf3-1 : ---------- : -
LcELF3 : MSPSEGLVYKPYPGPGFTGTNFGG-CGPYGASPSGGTFMNPSYGIPP--- : 534
PsELF3 : MSPSEGLVYKPYPGPGFTGTNFGG-CGPYAAAPSGGTFMNPSYGIPP--- : 534
MtELF3 : MSPSEGLVYKPYPGPGFTGTNYGG-SGPFGAPPSCGTFMNPSYGMPP--- : 520
CaELF3 : MSPSEGLVYKPYPGPGFTGTNCGE-YGPIGAAP----FMNPSYGMPA--- : 531
GmELF3 : MTPSEGLVYKPYPGPGFTGTGCGGGCGPFVPALLGGSFMNPGYGIPTSHQ : 539
AtELF3 : MSPSEGLIYKPHPGMAHTG-HYGGYYGHYMPTPMVMPQYHPGMGFPP--- : 521
560 * 580 * 600
Lcelf3-1 : ---------- : -
LcELF3 : ----PPEIPPGSHAYFPPYGGMPVMKAAASESAVEHVNQFSAH----GQN : 576
PsELF3 : ----PPETPPGSQAYFPPYGGMPVMKAAASESAVEHVNQFSAR----GQS : 576
MtELF3 : ----PPETPPGSHAYYPPYGSMPFMKAAASESVVEHVNQFSAR----VQS : 562
CaELF3 : ----PPETPPGSHAYFPPYGGMPVTKAAVAESAVGHVNQFSAH----GQN : 573
GmELF3 : GVGVPPDTHPGSHGYLPPYG-MPVMNSSMSESVVEQGNQFSALG-SHGHN : 587
AtELF3 : ----------PGNGYFPPYGMMPTIMNPYCSSQQQQQQQPNEQMNQFGHP : 561
* 620 * 640 *
Lcelf3-1 : ---------- : -
LcELF3 : HRLSED-EDNCNKHNQSSCNLPAQRNEDTSHVMYHQR-----SKEFDLQM : 620
PsELF3 : RRLSED-EADCNKHNQSSYDLPVQRNGATSHVMYHQR-----SKEFEVQM : 620
MtELF3 : RHLSEG-EADCNKHNQSSCNLPVQRNGATTHVMHHQR-----SKEFELQM : 606
CaELF3 : DHLSEV-EANHNKHNQIPCNLPAQRNGATSHVMNRQR-----HKEFELQG : 617
GmELF3 : GHLPGGGKANHNTNNKSSCNLPVQRNGAISHVLKHQT-----SKDFELQE : 632
AtELF3 : GNLQNTQQQQQRSDNEPAPQQQQQPTKSYPRARKSRQGSTGSSPSGPQGI : 611

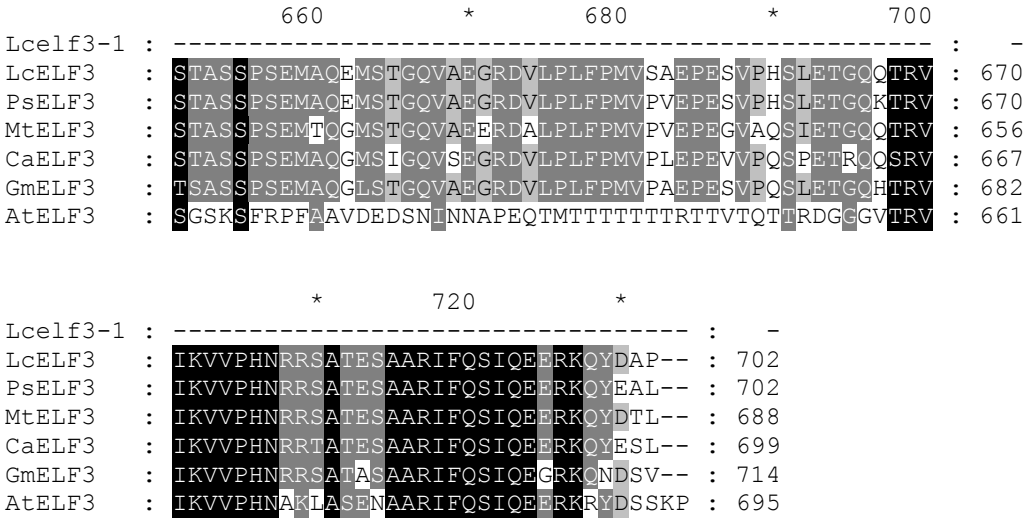


Figure 3-6 ELF3 predicted protein alignment.
The alignment was created with full-length ELF3 protein sequences of selected legumes and Arabidopsis aligned with ClustalX and manually adjusted and annotated using GeneDoc and Adobe Illustrator. Shading indicates degrees of conservation; black=100%, dark grey=80%, light grey=60%, green=conserved aspartic acid (D), red=exon 3 skipping. Species abbreviations are as follows: *Lens culinaris* (Lc), *P. sativum* (Ps), *M. truncatula* (Mt), *C. arietinum* (Ca), *G. max* (Gm), Prevalence of elf3-1 in a Lentil Association Mapping panel

3.3.6 Prevalence of elf3-1 allele in a Lentil Association Mapping panel

Understanding the prevalence of the early flowering *elf3-1* allele in an existing representative germplasm provides an insight into the origins of the early-flowering allele, and the role of the lentil *Sn* in flowering time adaption. A Lentil Association Mapping panel comprising of 94 accessions (Appendix 1) was analysed for the translationally silent G-to-A substitution. In this analysis, only ILL 4605, the early-flowering parent of ILL 6005, was identified to carry the substitution (Figure 3-7). Additionally, it was observed that the 18 *pilosae* and four *aethiopicae* accessions within the panel did not carry the early-flowering allele (refer to Appendix 1 for passport information). Individual accessions with ambiguous genotypic data were re-analysed (not visualised) to confirm absence of substitution.

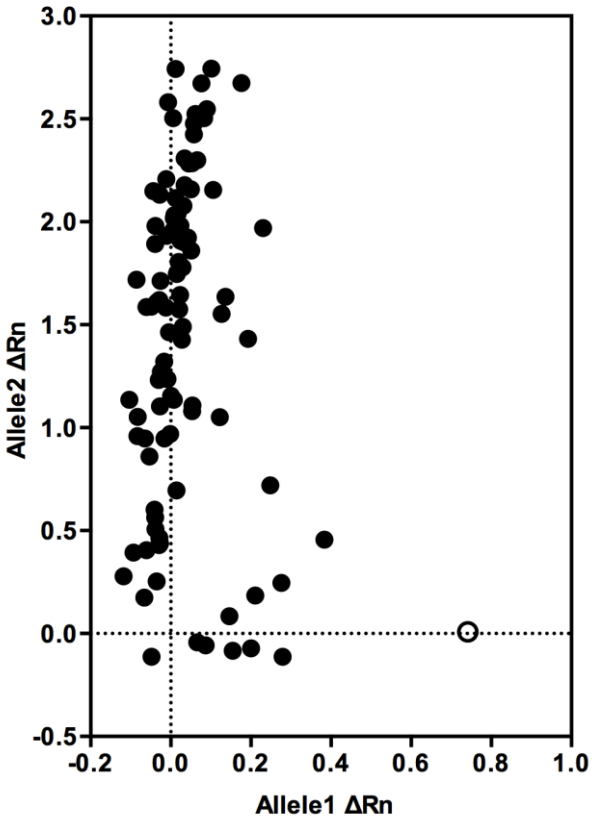


Figure 3-7 KASP assay of *elf3-1* prevalence across 94 lentil accessions.
Accessions genotyped using two lentil *ELF3*-specific forward primers and a single reverse primer. Closed symbols represent accessions carrying the functional *ELF3* allele and open symbol represent accessions carrying the early-flowering *elf3-1* allele.

3.3.7 Segregation of ILL 223 x ILL 5588 F_2 population

While the observed splicing defect is the most likely cause of the early-flowering phenotype, it is also possible that the second polymorphism involving the deletion of a single conserved aspartic acid residue may also contribute to the photoperiod-insensitivity and early-flowering phenotype observed in cv. Precoz and its derivatives.

To evaluate the possibility that the deletion might be functionally significant, available germplasms were surveyed for accessions that had the 3-bp deletion in exon 2 but not the translationally silent G-to-A substitution in the last nucleotide of exon 3. A mid-early flowering accession ILL 223, from East Azerbaijan in Iran was evaluated to carry the 3-bp deletion in exon 2. This allele from ILL 223 is designated *elf3-2* in this thesis.

ILL 223 was first evaluated for its photoperiod response under controlled SD and LD conditions. It was observed that the accession flowered significantly later ($p < 0.05$) in SD (65.6 ± 0.889 days) when compared to LD (30.1 ± 3.09 days), suggesting that ILL 223 is photoperiod-sensitive (Figure 3-8). ILL 223 was additionally observed to flower earlier than ILL 5588 in SD.

An F_2 population from a cross between ILL 223 and ILL 5588 was then established to evaluate for co-segregation between flowering time and *LcELF3*. A bimodal segregation distribution was observed in the F_2 progeny, with an early class and a late class (Figure 3-8). However, these discrete classes were observed not to co-segregate with *LcELF3* (Figure 3-8C). Progeny carrying the *elf3-2* allele were evaluated to not flower significantly (P = 0.104) earlier (Figure 3-8D) or at an earlier developmental stage (P = 0.338) (Figure 3-8E) when compared to progeny carrying the ILL 5588 allele. Therefore, while ILL 223 does flower earlier than ILL5588 in SD conditions, this difference cannot be attributed to *LcELF3*.

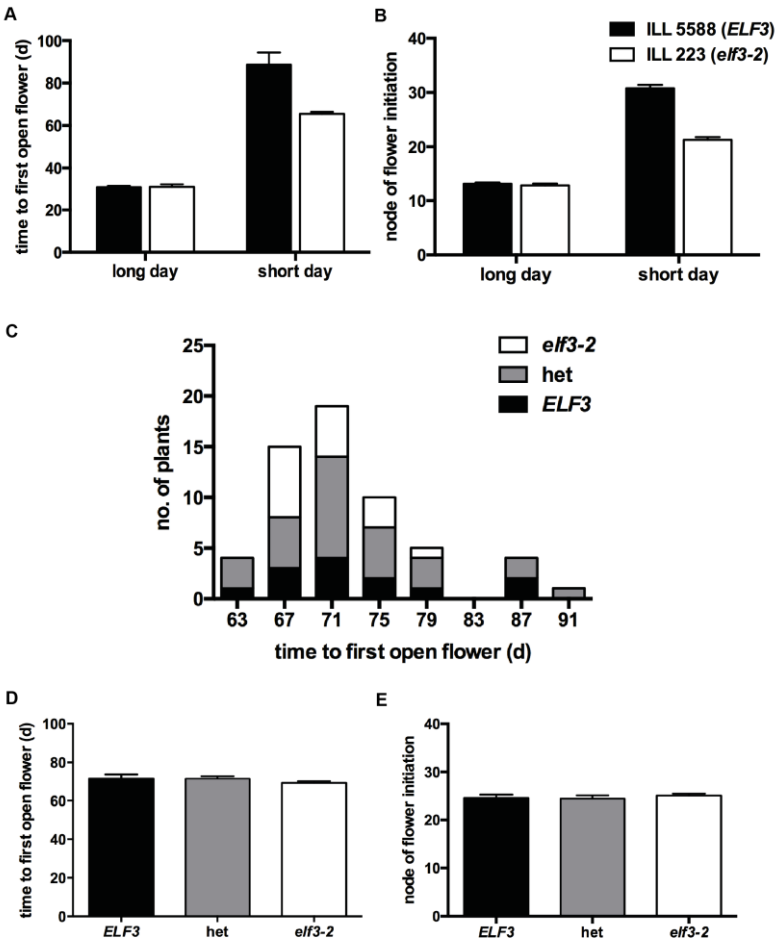


Figure 3-8 Phenotypic characterisation of *elf3-2*.
(A) Flowering time scored as days to first open flower in ILL 5588 and ILL 223 under LD and SD conditions. (B) Node of flower initiation, denoting the developmental age at the point of transition to reproductive development in ILL 5588 and ILL 223 under LD and SD conditions. Data for (A) and (B) are mean ±SE for *n*=5-8. (C) Co-segregation of early-flowering phenotype with *LcELF3* under SD conditions. (D) Association analysis of *LcELF3* with flowering time under SD conditions. (E) Association analysis of *LcELF3* with flowering node under SD conditions. Data for (D) and (E) are mean ±SE for *n*=13-29. Plants received a 12-h photoperiod of natural daylight (SD), extended with 4-h of florescent light (LD).

3.3.8 Effect of LcELF3 on other phenotypic traits

The current understanding of the relationship between the lentil *Sn* and other quantitative traits is limited. In *Arabidopsis*, *ELF3* has been previously implicated for its role in the regulation of hypocotyl elongation (Lu et al., 2012; Nusinow et al., 2011; Zagotta et al., 1996). In *P. sativum*, *ELF3* has been attributed to the propensity for branch formation (Weller et al., 2012).

In order to examine whether *ELF3* might also have broader pleiotropic roles in lentil, the effect of *ELF3* genotype on a range of other traits was also examined in the ILL5588 x ILL6005 F_2 population (Figure 3-9). These traits included *time to emergence* (DTE), *internode length between nodes 1 and 9* (IN9), *number of branches at 3 weeks* (EBN), and *total branch length at 3 weeks* (EBL).

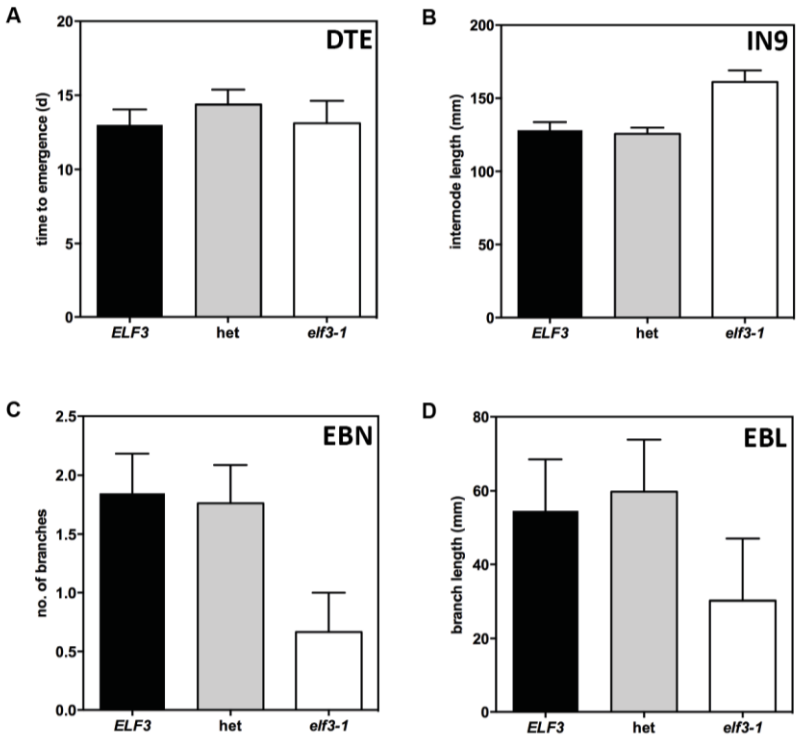


Figure 3-9 Association of *LcELF3* to other quantitative traits .
(A) Time to emergence from sowing. (B) Internode length between nodes 1 and 9. (C) Number of branches at 3 weeks from emergence. (D) Total branch length at 3 weeks from emergence.

Both IN9 and EBN show significant association with *LcELF3* genotype. F_2 progeny homozygous for the *elf3-1* were observed to have significantly elongated internodes ($p < 0.05$), and a significantly reduced propensity for branching ($p < 0.05$) when compared to progeny carrying the *ELF3* allele. No association between the *LcELF3* genotype was determined for DTE and EBL in the analysis of the homozygous classes ($p = 0.951$ and $p = 0.279$).

3.4 Discussion

This chapter determined the physiological and molecular basis for the lentil *Sn*, provided an insight into its control of key agronomic traits, and described the prevalence of the early-flowering allele across a representative collection.

3.4.1 Phenotypic characterisation of lentil Sn

This study first sought to establish the physiological basis for the early-flowering habit under controlled photoperiod conditions. Observations in section 3.3.1, determine that the lentil *Sn* confers ILL 6005, a cv. Precoz derivative, its early-flowering phenology by affording photoperiod-insensitivity (Figure 3-1). This is consistent with early physiological work by Summerfield et al. (1985) that described cv. Precoz as the least sensitive of studied accessions to photoperiod in inductive and non-inductive photo-thermal conditions.

In both *Arabidopsis* (Herrero et al., 2012; Kolmos et al., 2011; Liu et al., 2001; Lu et al., 2012; Matsushika et al., 2002; Mockler et al., 1999; Nusinow et al., 2011) and the phylogenetically related *P. sativum* (Hecht et al., 2007; Liew et al., 2009a; Liew et al., 2014; Weller et al., 2004), several photoreceptor and circadian clock mutants are described for their photoperiod-insensitive early-flowering phenotype. Comparative analogy of these mutants suggests that the lentil *Sn* is likely to function similarly to afford photoperiod-insensitivity.

The study defined the response of ILL 6005 to monochromatic light to determine if defects in its photoreceptors are indeed responsible for its flowering phenology. However, no well-defined photo-morphological defects as illustrated in Figure 3-2 were observed, therefore ruling out photoreceptors and their associated genes as potential candidates. This finding also confined further analysis of the lentil *Sn* to components of the circadian clock.

To evaluate if circadian clock components are responsible for the photoperiod response of ILL 6005, a F_2 population segregating for flowering time was established and evaluated under a short day photoperiod. A co-segregation

analysis for candidate genes involved in the circadian clock, as detailed in Table 3-3, was subsequently carried out. A complete co-segregation with the lentil orthologue for *Arabidopsis ELF3* was observed.

3.4.2 Molecular identity of lentil Sn

The co-segregation of the lentil *ELF3* orthologue to flowering time strongly suggests that the *LcELF3* is tightly linked, and is the most likely candidate for the lentil *Sn* locus (Figure 3-4). This is further supported by the identification of a substitution of the terminal nucleotide of *LcELF3* exon 3 that disrupts splicing of the *LcELF3* transcript and results in skipping of intron 3 (Figure 3-5A and Figure 3-6). This leads to a frame-shift mutation and a consequent premature stop codon, which results in the truncation of the predicted ELF3 protein (Figure 3-5A and Figure 3-6).

The dominant nature of inheritance and the role of the lentil *ELF3* in conferring an early-flowering phenology in its recessive state corresponds to previous work on the lentil *Sn* locus by Sarker et al. (1999), where it was proposed that the locus confers segregants with an early flowering phenology only in its recessive state. The late-flowering phenology of the progeny heterozygous for *LcELF3*, indistinguishable from the progeny homozygous for the ILL 5588 allele (Figure 3-4) further supports the bimodality of the segregation observed in previous work by Sarker et al. (1999).

Apart from the single substitution that results in a splicing defect, a 3-bp indel resulting in the deletion of a conserved aspartic acid was also noted in the mutant *elf3-1* allele. This polymorphism however, can be excluded as a significant contributor to the early-flowering phenology observed in cv. Precoz and its derivatives. In the segregation analysis between the earlier-flowering ILL 223, carrying the 3-bp deletion and ILL 5588, *LcELF3* was not observed to co-segregate for flowering time. In this analysis, it is important to recognise that ILL 223 displayed photoperiod sensitivity, and that while bimodality was observed

in the segregation for flowering time, it is likely that another major locus is responsible for the observed phenotype.

3.4.3 Role of ELF3 in circadian clock and regulation of flowering time

The circadian clock is an endogenous molecular feedback system that is characterised by interloping transcriptional regulators that enables plants to optimally respond to diurnal and rhythmic changes in the environment. The Arabidopsis *ELF3* is a member of the evening complex within this system which includes *LUX* and *ELF4* (Nusinow et al., 2011). *ELF3* has been extensively studied for its role in the maintenance of the circadian rhythm and its contribution to the regulation of flowering time (Amasino, 2010; Herrero et al., 2012; Kolmos et al., 2011; Nusinow et al., 2011).

ELF3 was first identified and reported in *Arabidopsis* mutants displaying photoperiod insensitivity and an early-flowering phenology in short days (Zagotta et al., 1996). The *Arabidopsis ELF3*, has also been associated with the regulation of vegetative photomorphogenesis, including hypocotyl elongation (Zagotta et al., 1996). More recently, the role of *ELF3* orthologues in the circadian clock, and in the control and regulation of flowering time has been construed in several crop plants including *P. sativum* (Weller et al., 2012), *H. vulgare* (Faure et al., 2012) and *O. sativa* (Matsubara et al., 2012). It has also been described in *P. sativum* that a mutation in its *ELF3* orthologue is responsible for the adaptation of the legume crop to spring-sowing, and its spread to the higher lattitudes (Weller et al., 2012). *ELF3* is also implicated in the control of flowering time in the ornamental sweet pea (*Lathyrus odoratus*) (Weller et al. unpublished), another phylogenetically related legume.

3.4.4 Pleiotropic effect of lentil Sn

Prior to this study, there was a limited understanding of the relationship between the lentil *Sn* and pleiotropic traits of agronomic significance. A strong association of the lentil *Sn* to internode length and early lateral branching was determined, with recessive *elf3-1* alleles conferring elongated internodes

(Figure 3-9B), and reduced propensity for early branching (Figure 3-9C). The ILL 5588 allele is dominant in both instances.

3.4.5 Contribution of lentil Sn to adaptation and spread

The genetic characterisation of the lentil *Sn* is a crucial first step towards developing our current understanding of flowering time control in lentil.

However, based on a survey for the prevalence of the *elf3-1* allele in a Lentil Association Mapping panel consisting of cultivated lentil accessions, it can be inferred that the lentil *Sn* was not responsible for the adaptation and spread of lentils from the Fertile Crescent and into the lower latitudinal production regions of India and Ethiopia. In this study, all 18 *pilosae* lentil accessions and all four *aethiopicae* lentil accessions were evaluated not to carry the early-flowering *elf3-1* allele (Figure 3-7).

This result is consistent with inferences of the *pilosae* lentil based on studies of lentil genetic diversity (Alo et al., 2011), and the continuous segregation reported by Sarker et al. (1999) in crosses between cv. Precoz and its derivatives, and early-flowering Indian accessions. The absence of the early-flowering *elf3-1* allele in all four *aethiopicae* lentil also present an expanded understanding of flowering time adaption in lentil.

3.4.6 Limitations of study

The unavailability of near-isogenic lines limited the scope of the project, and the capacity to present strong correlations of observed differences to *LcELF3*.

Chapter 4 Characterising the genetic control of earliness in an Indian landrace

4.1 Introduction

The *pilosae* lentil adapted to the agro-ecological environments of the Indian Subcontinent is morphologically distinct, and displays reduced genetic variability when compared to landraces from other regions (Cubero et al., 2009; Erskine et al., 1998; Erskine et al., 2011; Rana et al., 2007). Characterised by endemic traits that include pubescence on vegetative organs (Barulina, 1930), and the absence of tendrils (Vandenberg and Slinkard, 1989), the *pilosae* lentil also demonstrates an early flowering and an early maturing phenology that are attributed to its reduced photoperiod sensitivity and increased reliance on temperature cues for flowering (Erskine et al., 1990a; Erskine et al., 1994; Summerfield et al., 1985).

4.1.1 Origins of the pilosae lentil

It is hypothesised that cultivated lentil was first introduced into the Indo-Gangetic plain from West Asia through Afghanistan (Erskine et al., 2011). This is supported by Ferguson et al. (1998), which determined a close phylogenetic relationship between the *pilosae* lentil and the Afghan germplasm. Erskine et al. (1998) has put forward the notion that selection for local adaption coupled with the consequent reproductive isolation imposed by the asynchrony of flowering between West Asian germplasm and the *pilosae* lentil led to a genetic bottleneck. This genetic bottleneck is described to be responsible for the agro-morphological distinctness of the *pilosae* lentil (Erskine et al., 1998).

4.1.2 Flowering time and adaptation of the pilosae ecotype

Landraces adapted to the higher latitudes of Afghanistan represent some of the latest flowering lentil germplasms (Erskine et al., 1989; Erskine et al., 1990a; Erskine et al., 1994). The adaptation of this germplasm to the lower latitudes of

the Indo-Gangetic plain would have required the selection for an early flowering and early maturing phenology.

Erskine et al. (2011) proposes that the successful spread and adaptation of cultivated lentil into the Indo-Gangetic Plain is likely the result of repeated positive selection of recessive alleles for flowering time at intermediate elevations in Pakistan. The authors also raise the interesting possibility that these alleles could have arisen from the introgression with *L. culinaris* ssp. *orientalis* in Afghanistan.

4.1.3 Genetic basis for early-flowering in the pilosae lentil

Our current understanding of the early flowering habit of the *pilosae* ecotype is predominantly based on early physiological observations (Erskine et al., 1990a; Erskine et al., 1994; Summerfield et al., 1985). The existing literature presents limited information on the genetic and molecular basis for the flowering phenology of the *pilosae* lentil.

In the most significant genetic study on flowering time in lentil, Sarker et al. (1999) reported that F_2 progeny from crosses between cv. Precoz and early-flowering Indian accessions displayed a continuous segregation for flowering time, with early and late transgressive segregants. These observations led Sarker et al. (1999) to propose that a polygenic system shaped by interactions between the major lentil *Sn* locus and several minor loci was involved in the control of flowering time in lentil (Sarker et al., 1999). Chapter 3 determined that the lentil *Sn* is an *Arabidopsis ELF3* orthologue, and that Indian accessions did not carry the early-flowering *elf3-1* allele.

4.1.4 Chapter aims

This chapter aims to investigate the genetic basis for the earliness observed in ILL 2601, a *pilosae* accession that is evaluated to be amongst the earliest flowering in the lentil germplasm (Erskine, W. et al., pers. comm; Weller and Murfet, unpublished).

4.2 Materials and methods

This section details specific materials and methods relevant to this chapter. General materials and methods are described in Chapter 2.

4.2.1 Plant materials and growth conditions

A total of 173 F_2 individuals derived from a single cross between an early-flowering selection of the ILL 2601 Indian landrace, Line 24, and the photoperiod-sensitive ILL 5588 (cv. Northfield) were sown in February 2013 and evaluated under a base photoperiod of 12-hours of natural daylength at the University of Tasmania phytotron.

A 12-hour photoperiod was chosen, as preliminary experiments indicated sub-optimal growth under 10-hour short day photoperiod conditions, with photoperiod-sensitive ILL 5588 parental line exhibiting a stunted, abnormal appearance, accompanied by the failure to initiate flowers.

The parental accessions, ILL 2601 and ILL 5588, were evaluated under a base photoperiod of 12-hours of natural daylength for short days, and extended with 4-hours of fluorescent light for long days.

All seeds were scarified and imbibed for 12 hours in Milli-Q water prior to sowing.

4.2.2 DNA extraction

The genotyping technique employed for this study required that genomic DNA were subject to minimal shearing during extraction. To reduce shearing of genomic DNA during tissue lysis, tissue samples were firstly manually ground with a mortar and pestle. Genomic DNA was subsequently extracted using the protocol described in Chapter 2. All samples were assessed for digestibility using Taq$^{\alpha}$I before genotyping.

4.2.3 Diversity Array Technology (DArT) genotyping

The DArT-Seq™ (http://www.diversityarrays.com/dart-application-dartseq) high-throughput genotyping method was employed to generate Single Nucleotide Polymorphism (SNP) data for genetic analyses in this study. The genotyping was performed externally at Diversity Array Technology Pty. Ltd., Canberra, Australian Capital Territory.

The F_2 individuals were genotyped for a total of 9315 markers. Markers that were identified to be heterozygous for either parent, non-polymorphic, or had more than three F_2 individuals with missing data were excluded from subsequent analysis. A subset of 2161 polymorphic markers were retained and utilised in the construction of the genetic linkage map.

4.2.4 Construction of genetic linkage map

A genetic linkage map was constructed for the ILL 2601 x ILL 5588 F_2 population using JoinMap 4.0 (Van Ooijen, 2006). The independence logarithm of odds (LOD) significance threshold was utilised in a manner of increasing stringency to assign statistically associated polymorphic markers into groups. A minimum LOD value of 10.0 was used as the significance threshold to assign markers to groups using the *tree* command.

The regression algorithm was applied with the Kosambi function to resolve the order of the polymorphic markers and the distances between markers within each group. The regression mapping procedure estimates the relative position of each polymorphic marker by comparing the goodness-of-fit of the calculated map for each position (Van Ooijen, 2006). This procedure starts with the most informative pair of markers and is continued with the addition of every subsequent marker. Markers that result in negative distance estimates are excluded during this process. Once all the markers assigned to a specific group are tested for their positional estimates, a framework of markers for the group is formed, known as *first round* in the application (Van Ooijen, 2006). A *second round* is then undertaken to include the excluded markers using the pairwise

data into the existing framework. The application allows for an additional *third round* (Van Ooijen, 2006).

In order to construct a genetic linkage map with robust order, a goodness-of-fit threshold LOD of 5.0 was applied for the removal of loci in the first and second rounds of mapping. The *third round* function was not employed, instead after the first iteration of regression mapping, markers within ±1 cM of another marker were manually removed and another iteration of regression mapping was employed. The removal of markers was directed at reducing the number of markers in regions of the linkage map framework with high marker density. In subsequent iterations of regression mapping, markers demonstrating a high *nearest neighbour fit* (*N.N. Fit*), and a high *genotype probability* were progressively excluded from each linkage group until the target threshold for these parameters were achieved. Target thresholds of 6.0 *N.N. Fit* (cM) and 5.0 (-Log10(*P*)) were employed for these respective parameters.

In the first attempt of linkage map construction, only non-distorted ($p < 0.05$) markers, and markers that displayed reduced similarity (< 0.95) were included in the mapping procedure. In linkage groups where there were a larger number of excluded markers, using a threshold of 20% as a proportion of total makers, a secondary attempt at map construction was undertaken using the above-mentioned mapping procedure with all markers.

A set of 734 polymorphic markers was used in the construction of the final genetic linkage map. The linkage map for the segregating population was graphed using MapChart 2.3.

4.2.5 Medicago truncatula synteny

Synteny and collinearity between *L. culinaris* and *M. truncatula* was established by comparing the similarity of the nucleotide sequences of the final set of markers against the genomic sequences of the Mt4.0 version of the *Medicago* reference genome. Sequences demonstrating regions of significant local similarities (*e*-value < 0.001) were employed to ascertain the syntenic relationship of the final constructed linkage groups in *L. culinaris* with that of *M. truncatula*. This relationship was expressed using a dot plot, graphed using Prism 6 and visualised with Adobe Illustrator CS5. The left most position of the reference sequence was utilised as the reference position for each marker. An additional 38 sequences (*e*-value < 0.01) and 33 sequences (*e*-value < 0.1) demonstrating low similarity were included in the final dot plot.

4.2.6 Quantitative trait loci (QTL) analysis

QTL analysis was undertaken using MapQTL 6 (Van Ooijen, 2009). A primary analysis for each trait was carried out using the *Interval Mapping* function using phenotypic data from the 173 ILL 2601 x ILL 5588 F_2 individuals against the genetic linkage map. The QTL LOD significance threshold for each trait was determined using the *Permutation Test* function, where 1000 permutations at a significance level of $p < 0.05$ was employed. A *genome-wide* and *chromosome-wide* QTL LOD significance threshold score derived from the *Permutation Test* was used to determine QTL for traits scored in the population.

A secondary round of QTL analysis was then undertaken for traits with known QTL determined from *Interval Mapping*. The *Multiple-QTL Model* (*MQM*) function was employed in the secondary round of QTL analysis. This function was employed in conjunction with cofactors to reduce the residual variance attributable to a known QTL and to increase the power of the QTL analysis to allow the resolution of other segregating QTL for each studied trait. The *MQM* function was employed reiteratively with each new cofactor selection until all QTL for the observed variation for a specific trait were determined. A new QTL

LOD significance threshold was determined using the *Permutation Test* function with the selection of each new cofactor.

Cofactors were initially user-determined and then subject to a likelihood analysis based on backward elimination ($p < 0.05$) employed by the *Automatic Cofactor Selection* (*ACS*) function to determine their suitability for *MQM* analysis. This was carried out so as to reduce the incidence of a false secondary QTL (type I error) in the presence of a redundant cofactor.

4.2.7 Plant measurements

Refer to Chapter 2.

4.3 Results

4.3.1 Phenotypic characterisation of ILL 2601 under different photoperiods

The early habit of ILL 2601 was evaluated for its photoperiod-sensitivity under controlled long day (LD) and short day (SD) photoperiods. Observations were made relative to the photoperiod sensitive, medium-late flowering accession ILL 5588 (refer to 4.2.1 for photoperiod conditions).

Flowering time observations are often recorded as *days to flowering* (DTF) and *node of flower development* (NFD). In this study it was also observed that the node bearing the first open flower was not always the node at which the first floral structure was initiated. Instead in some plants, as illustrated in Figure 4-1A, nodes bearing aborted or undeveloped flowers preceded the appearance of a fully developed open flower. This phenomenon was anecdotally observed in ILL 5588 under SD in Chapter 3, and has been described in *Pisum sativum*, a phylogenetically related *Fabeae* legume (Berry and Aitken, 1979). To further explore the genetic basis for the flowering node, this chapter dissected NFD into two independent traits, namely *node of floral initiation* (NFI) which accounts for earliest node bearing aborted or undeveloped flowers, and *delay to flower development* (DFD) which accounts for the node interval between NFI and NFD (Figure 4-1A).

The period of the pre-emergent phase, first described by Roberts et al. (1986), is often included in the reported DTF (Erskine et al., 1990a; Saha et al., 2013; Tullu et al., 2008). Preliminary evidence point to a variation in the period for this phase across accessions, and its inclusion can distort the described variation for DTF. To afford a better understanding of this phase, defined in this thesis as the interval (time) between sowing and the appearance of the first pair of bifoliate leaves (Figure 4-1A), this phase was analysed as an independent trait and designated *days to emergence* (DTE).

In this chapter, ILL 2601 and ILL 5588 were characterised with respect to DTF, NFD, NFI, DFD, and DTE.

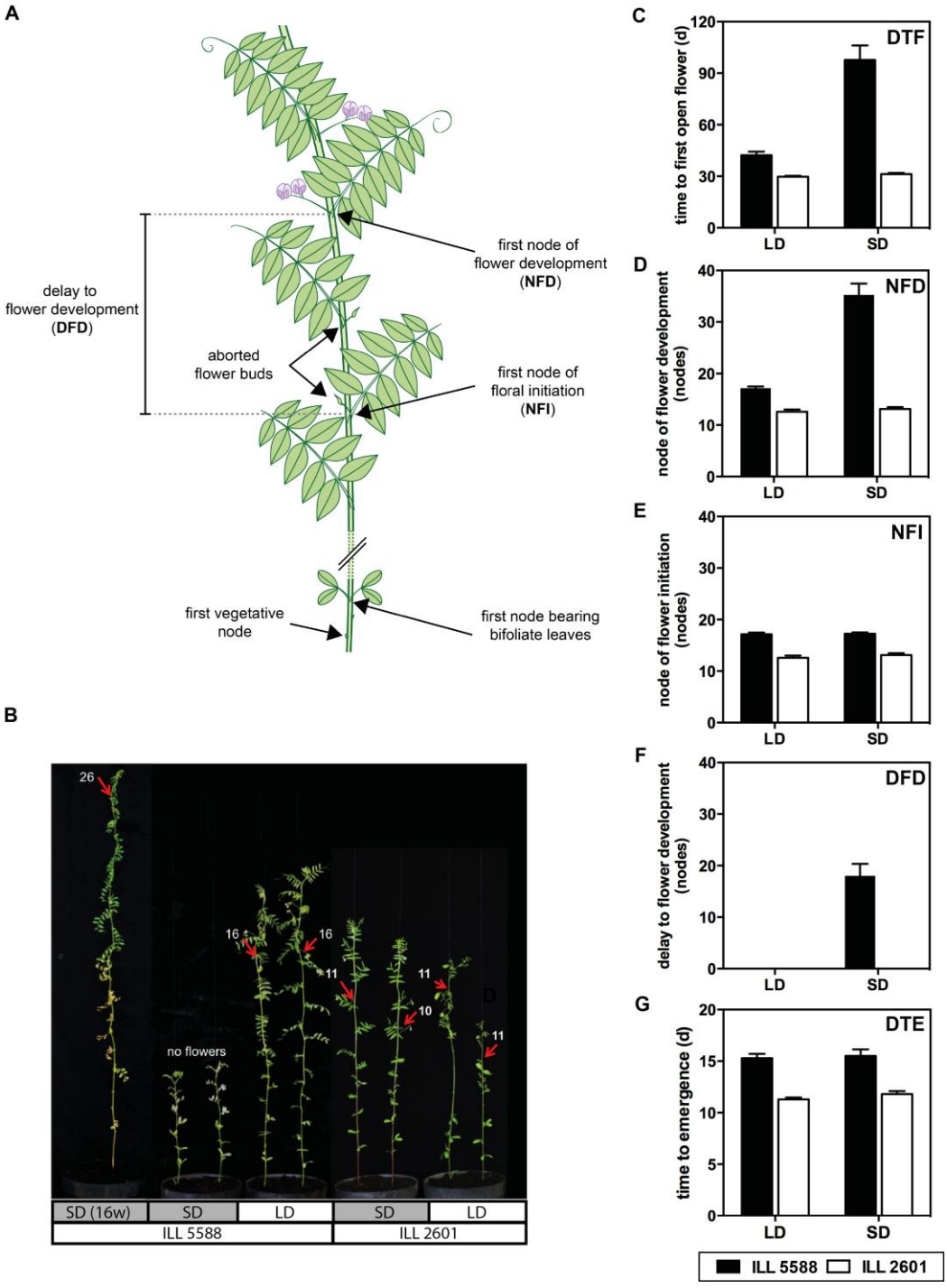


Figure 4-1 Phenotypic characterisation of ILL 2601 under different photoperiods
(A) Schematic of lentil plant with NFD, NFI, and DFD illustrated. (B) Representatives of ILL 5588 and ILL 2601 (lateral branches excised) grown under LD and SD conditions at 7-weeks from emergence, and ILL 5588 (with lateral branches excised) grown under SD at 16-weeks from emergence. Red arrows and numbers denote NFD. (C) *Days to flowering* (DTF) in ILL 5588 and ILL 2601. (D) *Node of flower development* (NFD), denoting the developmental stage of first developed/open flower in ILL 5588 and ILL 2601. (E) *Node of flower initiation* (NFI), denoting node bearing first floral structure in ILL 5588 and ILL 2601 (F) *Delay to flower development* (DFD), denoting node interval between NFI and NFD in ILL 5588 and ILL 2601. (G) *Days to emergence* (DTE), denoting period between sowing and appearance of first pair of bifoliate leaves. Plants received a 12-h photoperiod of natural daylight (SD) and a 12-h photoperiod of natural daylight extended with 4-hours of fluorescent light (LD). Data are mean ±SE for *n*=4-10.

The phenotypic characterisation of ILL 2601 determined that the landrace was not responsive to photoperiod, with no significant difference observed for DTF, NFD, NFI, and DFD (Figure 4-1C to Figure 4-1F) between LD and SD photoperiods. ILL 2601 was also observed to flower earlier than ILL 5588 for all measures (Figure 4-1). As expected, ILL 5588 was observed to be photoperiod sensitive for DTF, NFD, and DFD (Figure 4-1C to Figure 4-1F).

However, ILL 5588 was observed to not demonstrate a photoperiod response for NFI (Figure 4-1E), with no significant difference observed between LD and SD photoperiods. The observations imply that NFI is regulated independently of prevailing photoperiod. This is in contrast to observations for DFD that suggest a strong photoperiodic basis for the development of an open flower.

ILL 2601 seedlings were also observed to emerge earlier than ILL 5588. The pre-emergent phase is not responsive to photoperiod, with no significant difference observed between LD and SD photoperiods (Figure 4-1G). The reported photoperiod-insensitivity of this phase is consistent with early physiological observations by Roberts et al. (1986).

Observations for ILL 2601 and ILL 5588 are summarised in Table 4-1.

		LD		SD		*p*-value	Figure
		mean	± SE	mean	± SE		
ILL 2601	**DTF** (days)	29.7	0.522	31.2	0.757	0.1620	Figure 4-1C
	NFD (nodes)	12.6	0.429	13.1	0.379	0.3747	Figure 4-1D
	NFI (nodes)	12.6	0.429	13.1	0.379	0.3747	Figure 4-1E
	DFD (nodes)	0.00	0.00	0.00	0.00	N/A	Figure 4-1F
	DTE (days)	11.3	0.184	11.8	0.291	0.1986	Figure 4-1G
ILL 5588	**DTF** (days)	42.3	2.04	97.8	8.41	0.0000*	Figure 4-1C
	NFD (nodes)	17.1	0.340	35.3	2.21	0.0000*	Figure 4-1D
	NFI (nodes)	17.1	0.340	17.3	0.250	0.8334	Figure 4-1E
	DFD (nodes)	0.00	0.00	18.0	2.35	0.0000*	Figure 4-1F
	DTE (days)	15.3	0.421	15.5	0.645	0.7777	Figure 4-1G

Table 4-1 Summary of DTF, NFD, NFI and DFD under different photoperiods
Days to flowering (DTF), node of flower development (NFD), node of flower initiation (NFI), delay to flower development (DFD), and days to emergence (DTE) in ILL 5588 and ILL 2601 under LD and SD conditions. Plants received a 12-h photoperiod of natural daylight (SD) and a 12-h photoperiod of natural daylight extended with 4-hours of fluorescent light (LD). Asterisk (*) indicates statistical significance ($p < 0.05$). Data are mean ±SE for n=4-10.

4.3.2 Role of lentil Sn in ILL 2601

The early-flowering phenotype of ILL 2601 is consistent with flowering time observations of Indian accessions reported in the literature (Erskine et al., 1990a; Erskine et al., 1994), where reduced photoperiod-sensitivity is described. The flowering phenology of ILL 2601 is also similar to the photoperiod-insensitive early-flowering phenotype conferred by the early-flowering *elf3-1* alleles derived from cv. Precoz described in Chapter 3 (Figure 3-1).

Findings from Chapter 3 (Figure 3-7) however suggest that it is unlikely that the early-flowering phenology of ILL 2601 is conferred by the recessive *elf3-1* alleles. This proposes that in ILL 2601, the lentil *Sn*/*ELF3* gene should be intact and fully functional. The lentil *Sn*/*ELF3* gene in ILL 2601 was sequenced and its predicted protein analysed (Figure 4-2). No deleterious polymorphisms were observed in the coding sequence, excluding the role of the lentil *Sn* in conferring ILL 2601 its early-flowering phenotype (Figure 4-2).

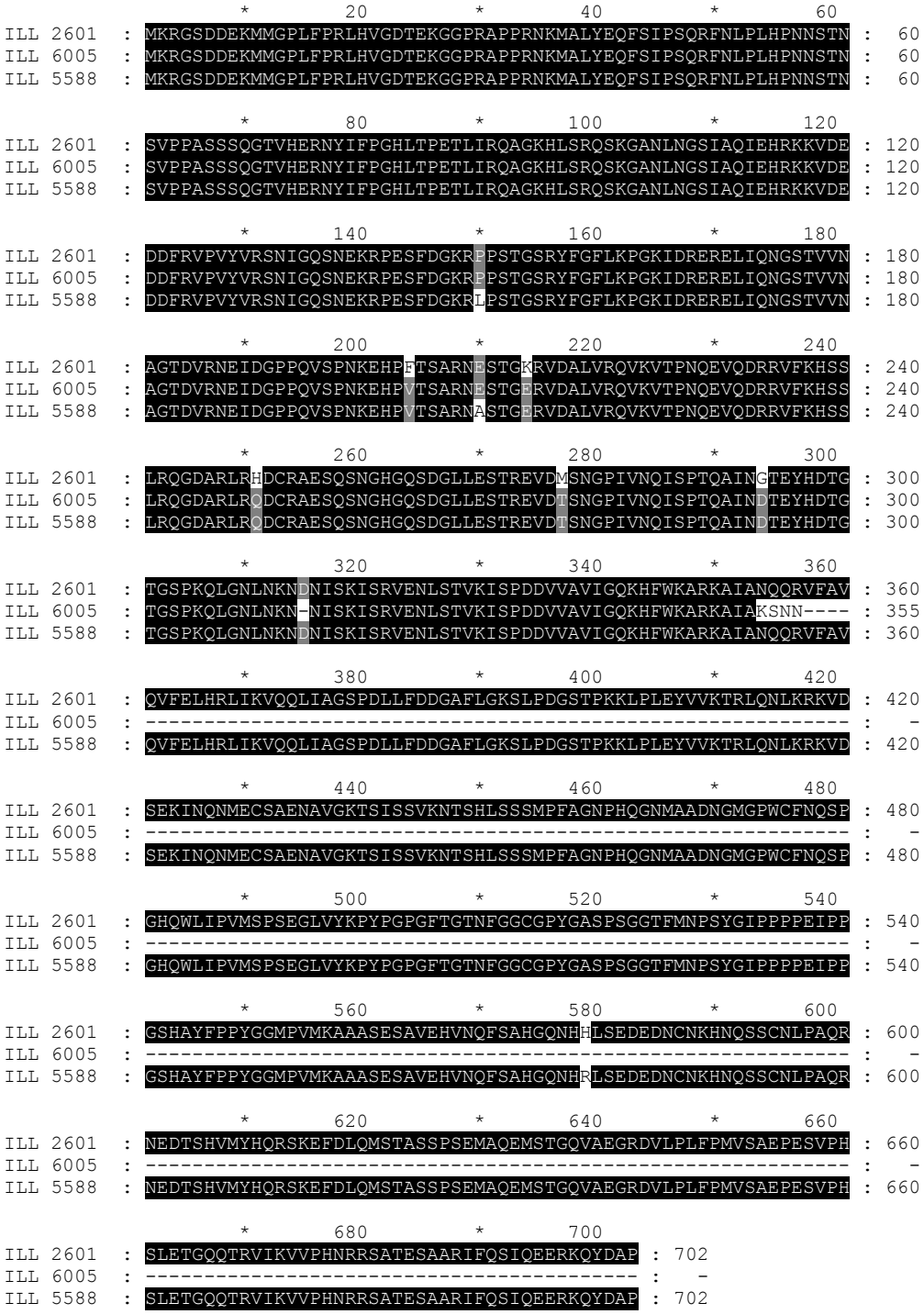


Figure 4-2 ELF3 predicted protein alignment.

The alignment was created with full-length predicted protein sequences from ILL 5588 (*Sn*/*ELF3*), ILL 6005 (*sn*/*elf3-1*), and ILL 2601 aligned with ClustalX and manually adjusted and annotated using GeneDoc and Adobe Illustrator. Shading indicates degrees of conservation; black=100%, dark grey=80%, light grey=60%. Refer to Appendix 4 for sequence details.

4.3.3 Segregation of ILL 2601 x ILL 5588 F_2 population for flowering time

To evaluate the genetic basis for the photoperiod-insensitive early-flowering phenology of ILL 2601, a F_2 population generated from a cross between ILL 2601 and photoperiod-sensitive ILL 5588 was established, and F_2 progeny evaluated under a 12-hour photoperiod (SD) of natural daylight.

A continuous distribution was observed for DTF in the ILL 2601 x ILL 5588 F_2 population (Figure 4-3A). It was also observed that the early segregants flowered at an earlier developmental stage with the initiation of the first open flower at a lower node. A strong positive correlation was determined between both DTF and NFD (R^2_{adj} = 0.905) (Figure 4-3B). Additionally, transgressive early and late segregants were observed in the ILL 2601 x ILL 5588 F_2 population, suggesting that minor loci are potentially involved in the control of DTF and NFD.

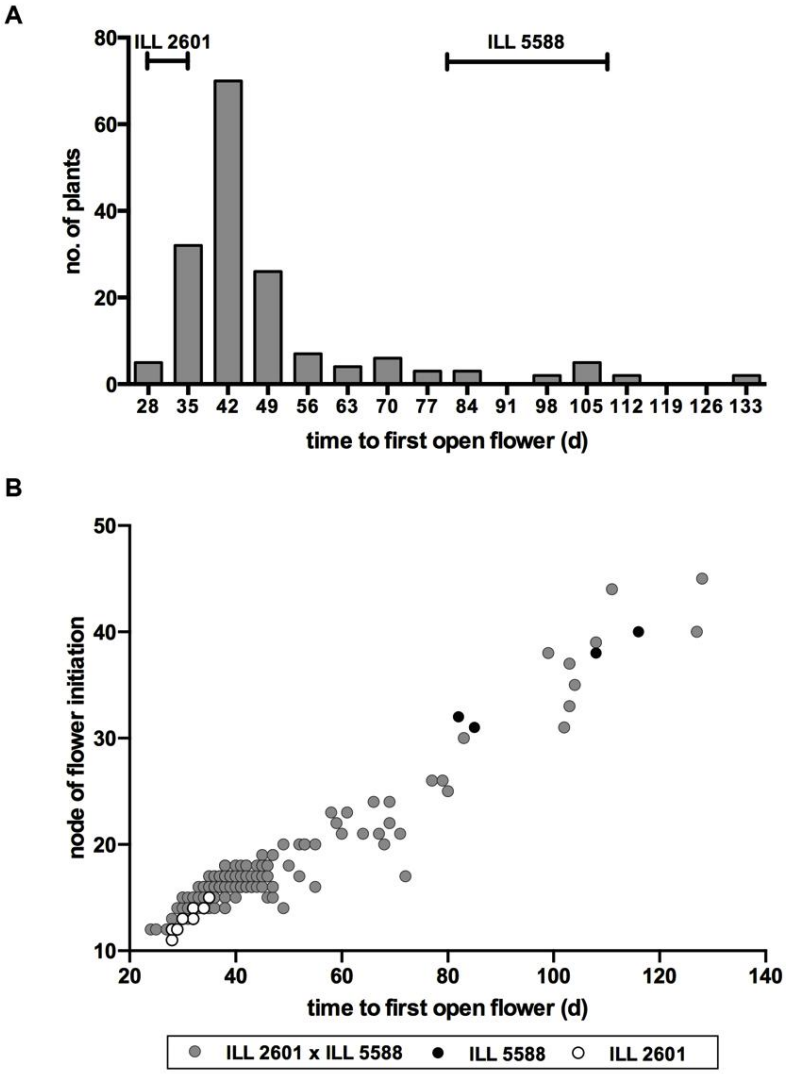


Figure 4-3 Segregation of ILL 2601 x ILL 5588 F_2 population for flowering time.
(A) Continuous distribution of F_2 progeny with respect to DTF under SD conditions. Data are for *n*=167. (B) Transition to reproductive development in F_2 progeny illustrated in the context of DTF and NFD under SD conditions. Data are for *n*=165. Plants received a 12-h photoperiod of natural daylight.

4.3.4 Genetic linkage map construction for ILL 2601 x ILL 5588 F_2 population

A genetic linkage map was constructed (refer to Section 4.2.4 for details) using DArT-Seq™ markers to determine the genetic basis for the observed variation in flowering time in the ILL 2601 x ILL 5588 F_2 population. While mapping populations have been extensively used in the study of phenotypic variation for key agronomic traits in lentil, a mapping population for a cross between ILL 2601 and ILL 5588 has not been previously reported.

The final genetic linkage map for the ILL 2601 x ILL 5588 F_2 population has an overall map length of 1032 cM defined by seven linkage groups, and 734 DArT-Seq™ markers (Figure 4-4). The average interval between the each segregating marker pair was 1.41 cM, with only one interval greater than 10 cM (Figure 4-4 and Table 4-2). The resolution of seven linkage groups corresponds to the seven chromosomes of the *Lens* genus (2n = 14).

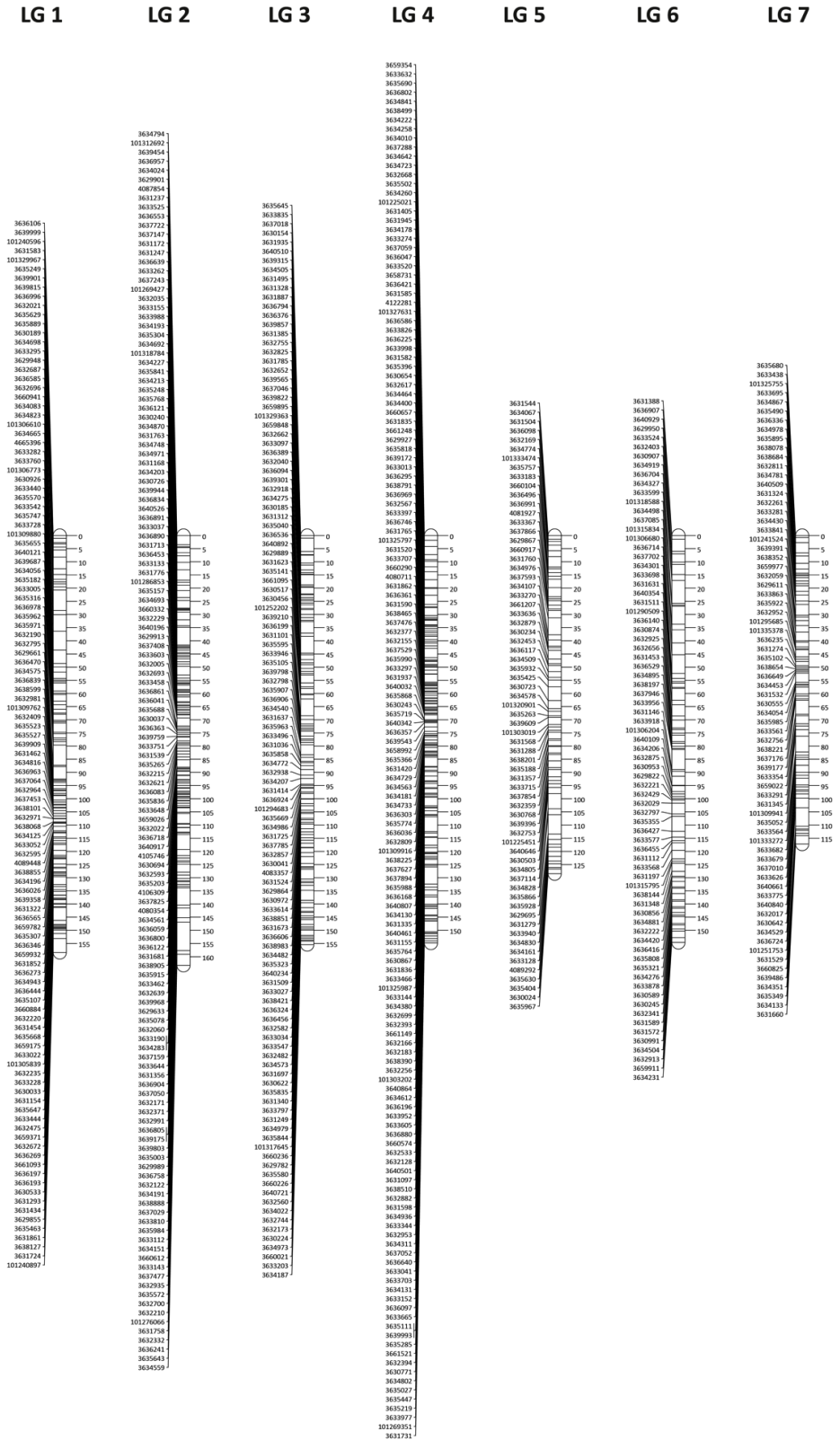


Figure 4-4 ILL 5588 x ILL 2601 F_2 genetic linkage map
Genetic linkage map consists of seven linkage groups corresponding to the seven chromosomes of the *Lens* genus. The nomenclature proposed is adapted from Sharpe et al. (2013).

The length of each linkage group ranged between 117 cM and 163 cM (Table 4-1). Different marker densities were also reported in the final linkage groups (Table 4-2). In the final data set used for map construction, 416 (19.3%) displayed significant segregation distortion ($p < 0.05$), the reported markers with segregation distortion within each linkage group was disparate (Table 4-2).

		Linkage Groups							
		1	**2***	**3**	**4***	**5***	**6**	**7**	**Total**
Primary Analysis	**Polymorphic**	269	345	368	361	276	299	243	**2161**
	Distorted (P<0.05)	30	113	57	69	91	23	33	**416**
	Similarity (>95%)	5	8	11	7	9	7	6	**53**
	Excluded (Primary)	35	121	68	76	100	30	39	**469**
Final Map	**Total markers**	115	136	118	151	67	75	72	**734**
	Total distance (cM)	158.40	163.35	155.43	154.81	128.51	154.63	117.16	**1032.28**
	Density (marker/cM)	1.38	1.20	1.32	1.03	1.92	2.06	1.63	**1.41**
	Intervals >10cM	0	1	0	0	0	0	0	**1**

Table 4-2 Linkage map summary
DArT-Seq™ markers categorised into the seven linkage groups as determined by a minimum LOD value of 10. *Polymorphic* denotes the total number of markers polymorphic for each loci, where the total number of missing data is ≤3. *Distorted* denotes the number of markers that demonstrated segregation distortion ($p < 0.05$). *Similarity* denotes the number of markers where > 95% of individuals report the same genotype as another marker. *Total distance* denotes the length of each linkage group. *Density* denotes the number of markers per cM. *Intervals > 10 cM* denotes the number markers where the distance between intervals exceeds 10 cM. Linkage groups 2, 4, and 5, indicated by (*) were determined using all markers.

A high proportion of markers demonstrating segregation distortion ($p < 0.05$) were detected in Linkage Groups 2, 4, and 5 (Table 4-2). The distorted markers, and markers demonstrating increased similarity (< 0.95) were included in the secondary mapping procedure for these linkage groups to avoid the exclusion of large contiguous regions of high distortion. In Linkage Groups 1, 3, 6, and 7, distorted ($p < 0.05$) markers, and markers that displayed increased similarity (< 0.95) were excluded in the mapping procedure.

4.3.5 Syntenic relationship between Medicago truncatula and Lens culinaris

A close macrosyntenic relationship between *L. culinaris* and *M*. *truncatula* has been previously reported (Ellwood et al., 2008; Kaur et al., 2014; Phan et al., 2006; Sharpe et al., 2013). The more recent work by Sharpe et al. (2013), presents the most comprehensive analysis of synteny between lentil and *M. truncatula, which* has considerably progressed our understanding of this relationship.

The sequences of the DArT-Seq™ markers incorporated into the final framework of the ILL 2601 x ILL 5588 genetic linkage map were analysed for regions of significant local similarities (*e*-value < 0.001) against the Mt4.0 version (Tang et al., 2014) of the *Medicago* reference genome to ascertain the syntenic relationship of the final constructed linkage groups in *L. culinaris* with that of *M. truncatula*, and improve on our current understanding of this relationship.

As illustrated in Figure 4-5, macrosynteny was established between the seven linkage groups for *L. culinaris* (2n=14) defined in the ILL 2601 x ILL 5588 genetic linkage map with that of the eight chromosomes of *M*. *truncatula* (2n=16).

The analysis also point to collinearity of *M*. *truncatula* chromosome 6 to the middle of the lentil linkage group 2, and translocations between the ends of the lentil linkage groups 1 and 5 (Figure 4-5). Major inversions were also observed in regions of lentil linkage groups 1 and 7 when compared to *M*. *truncatula* chromosomes 1 and 8. These reported findings are consistent with observations by Sharpe et al. (2013).

Additionally, translocations between the ends of the reported lentil linkage groups 4 and 8 was also reported in the ILL 2601 x ILL 5588 genetic linkage map. This is likely attributable to an aberrant chromosomal arrangement that resulted from a translocation event in the *M. truncatula* model accession A17, where the reciprocal translocation of the long arms of chromosomes 4 and 8

has been noted (Kamphuis et al., 2007). This reciprocal translocation however, was not presented by Sharpe et al. (2013). Separately, a large inversion in the lentil linkage group 3 reported by Sharpe et al. (2013) was also not reported in the ILL 2601 x ILL 5588 genetic linkage map.

Nevertheless, there was a general consensus between the ILL 2601 x ILL 5588 linkage map and that of Sharpe et al. (2013). Hence, the nomenclature for linkage groups used by these authors was adopted, and will be used hereafter in this study.

Comparative mapping of the lentil linkage groups with *M*. *truncatula* also allows for the estimation of the overall coverage of the developed ILL 2601 x ILL 5588 genetic linkage map, as a proportion of the annotated base pairs in Mt4.0. Briefly, based on the similarity of the nucleotide sequences of the DArT-Seq^{TM} markers on either end of the lentil linkage groups, an estimated gross coverage no less than 98.0% was derived for each of the *M*. *truncatula* chromosomes, with the exception of *M*. *truncatula* chromosome 6 where a 72.0% coverage was reported. However, there were large regions from *M*. *truncatula* chromosome 4, 5, and 7 that were not accounted for in the ILL 2601 x ILL 5588 lentil genetic linkage map.

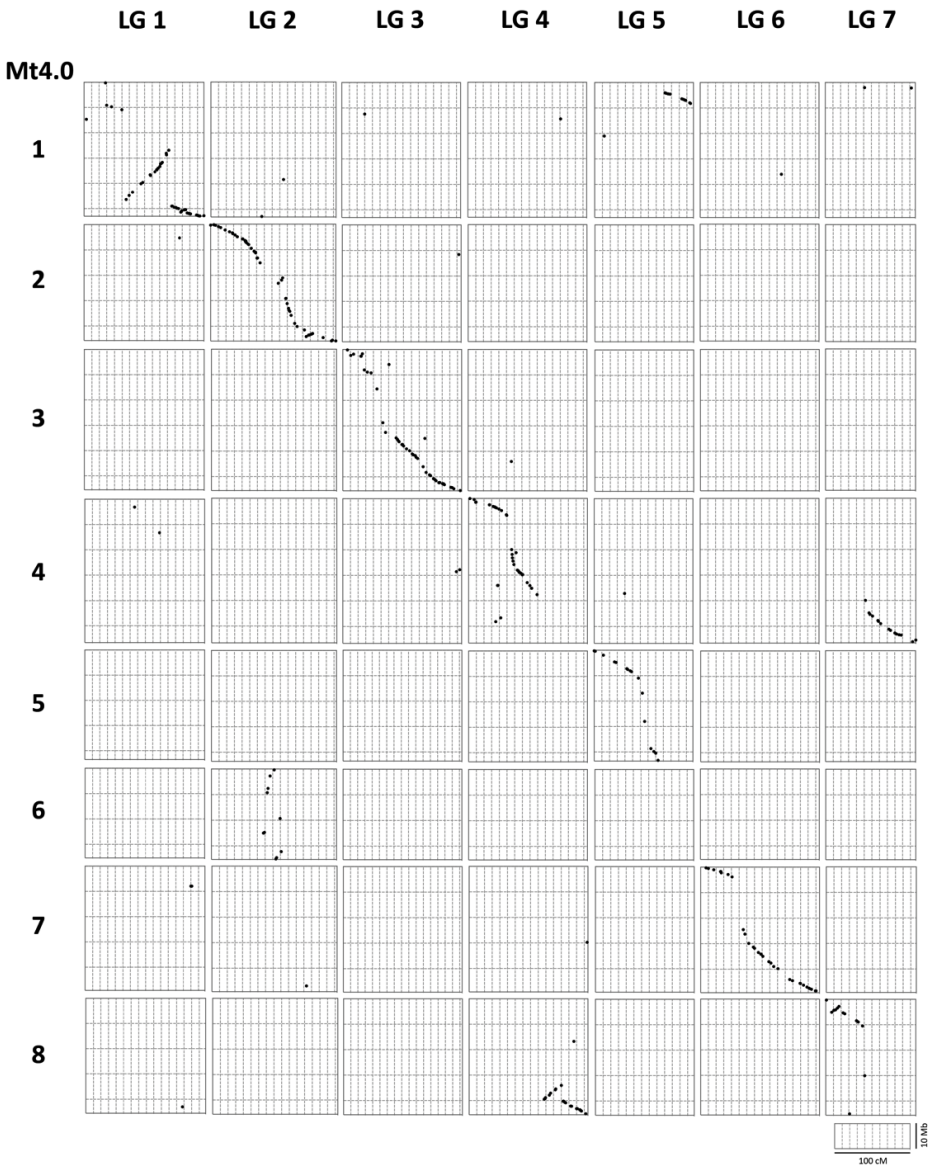


Figure 4-5 Dot plot of synteny between lentil and *M. truncatula* genome (Mt4.0)
The Y-axis represents the base pair position of the reference sequence of the Mt4.0 final assembly. The X-axis represents the distance between the mapped markers for each lentil genetic linkage groups. Each Y-axis interval represents 10 Mb and each X-axis interval represents 10 cM. 256 markers with sequences demonstrating significant similarities (*e*-value < 0.001), and 33 markers (*e*-value < 0.1) and 38 markers (*e*-value < 0.01) with sequences demonstrating low similarities are visualised in the dot plot. The dot plot was manually graphed using Prism 6.

4.3.6 Loci contributing to earliness of ILL 2601

The genetic control of the early-flowering phenology that characterises the *pilosae* lentil is poorly understood. While it has been suggested that the selection for recessive alleles for flowering time at intermediate elevations would have allowed for the adaptation of the cultivated species from Afghanistan to the Indo-Gangetic Plain (Erskine et al., 2011), and that a polygenic system is likely responsible for the early-flowering phenology of the *pilosae* lentil (Sarker et al., 1999), no QTL responsible for this adaptive shift have been proposed.

Through QTL mapping, four different qualitative traits were identified to collectively contribute to the observed variation for earliness in the ILL 2601 x ILL 5588 F_2 population. Two loci regulating DTF, and two loci regulating the *days to emergence* (DTE) from sowing were identified to be associated with the earliness observed in ILL 2601, when compared to ILL 5588 (Table 4-3). Three loci were also determined to contribute to the regulation of the complex trait NFD (Table 4-3). One of these was identified to be involved in the regulation of the NFI, denoting the appearance of the first floral structure, and two loci were determined to contribute to the delay between NFI and NFD, designated DFD (Table 4-3).

This section individually explores the contribution of each of these traits, and their associated loci to the observed earliness of ILL 2601, relative to ILL 5588 under SD conditions. A summary of loci determined through QTL mapping to contribute to earliness is detailed in Table 4-3.

	QTL information						Peak Marker Information		
Trait	**QTL**	**LG**	**Position (cM)**	**LOD threshold**	**Max LOD**	**% Expl. Variation**	**DArT-Seq™ Marker ID**	**Marker position (cM)**	**Marker LOD**
Days to flowering (DTF)	*DTF1*	6	67.3	5.5	17.6	33.9	101290509	68.6	16.7
	DTF2	6	152.3	5.5	7.1	12.0	3659911	152.8	7.0
Node of flower development (NFD)	*NFD1*	6	67.3	5.3	13.1	23.4	101290509	68.6	12.3
	NFD2	6	152.3	5.3	8.7	15.0	3659911	152.8	14.9
	NFD3	2	68.1	4.6	6.4	10.5	3632005	68.1	6.4
Node of floral initiation (NFI)	*NFI1*	6	67.3	5.1	8.0	20.0	3631511	64.3	6.8
Delay to flower development (DFD)	*DFD1*	6	152.8	5.6	11.0	14.6	3659911	152.8	11.0
	DFD2	2	69.1	5.8	6.0	7.4	3632005	68.1	6.0
Days to emergence (DTE)	*DTE1*	7	53.3	4.9	34.5	52.8	3631532	52.3	33.4
	DTE2	5	68.5	5.2	11.5	12.6	3635263	68.5	11.5

Table 4-3 Traits and QTL contributing to earliness in ILL 2601
Four quantitative traits and a complex trait; *days to flowering* (DTF), *node of flower development* (NFD), *node of floral initiation* (NFI), *delay to flower development* (DFD), and *days to emergence* (DTE) contribute to earliness. Max LOD refers to maximum LOD score for each trait determined using MQM. % Expl. Variation refers to percentage of total variation for a particular trait attributable to a QTL, determined using MQM. Marker LOD refers to LOD score for peak markers for specific traits determined using MQM.

4.3.6.1 Loci contributing to the variation in flowering time

A QTL analysis of DTF data from the segregating ILL 2601 x ILL 5588 F_2 population, evaluated under SD conditions, determined two loci for the observed variation in DTF. Both loci were in lentil linkage group 6 (Figure 4-6). For the purpose of this thesis, the two DTF loci have been provisionally assigned *DTF1* and *DTF2*, in the order of their contribution to the observed variation. *DTF1* and *DTF2* reported a maximum LOD score of 17.6 and 7.10, respectively, and accounted for an estimated 33.9% and 12.0% of the observed variation respectively (Table 4-2). In both instances, progeny homozygous for the ILL 2601 allele of the peak marker flowered earlier.

A chromosome-wide *LOD threshold* of 5.50 was employed for the determination of the QTL. The genome-wide *LOD threshold* reported was 7.70.

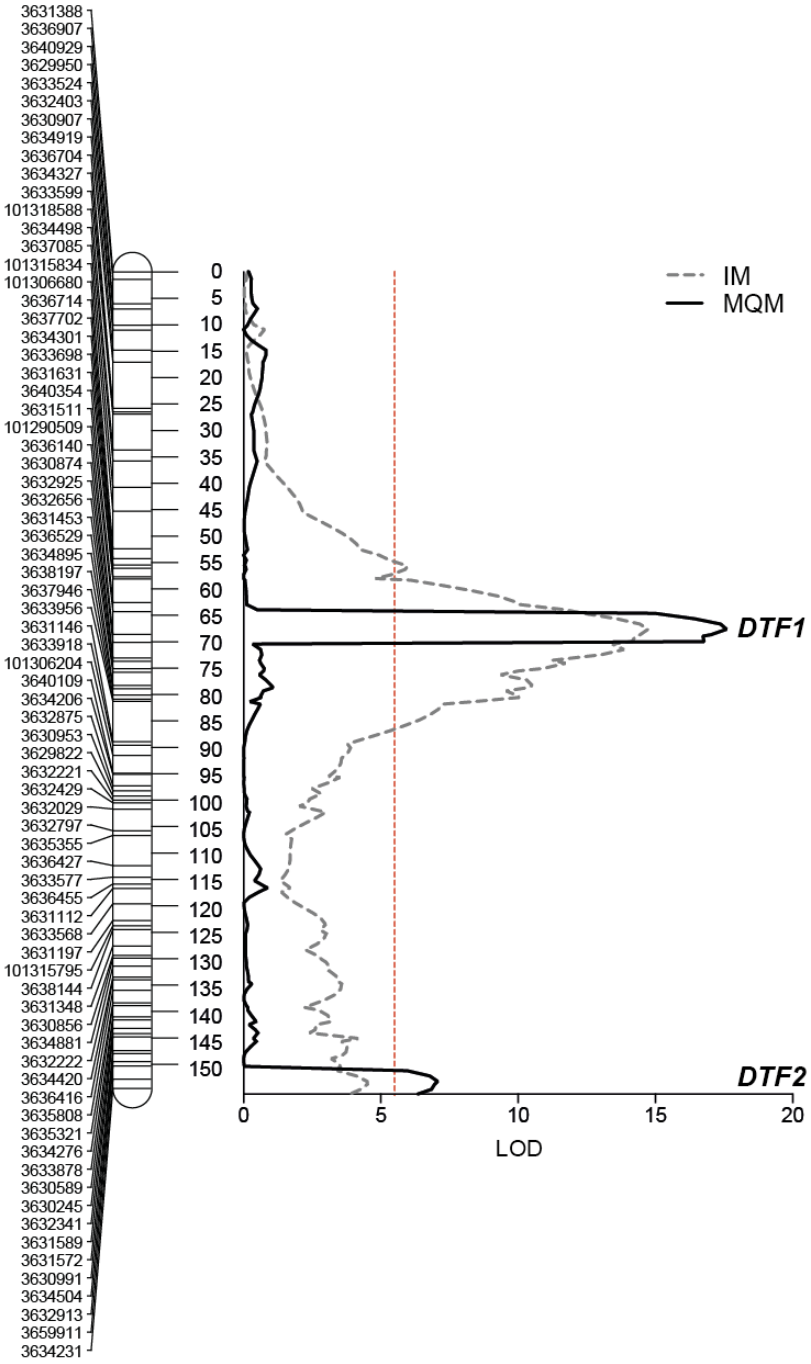


Figure 4-6 Flowering time loci in ILL 2601 x ILL 5588 F_2 population

Two loci, located on lentil linkage group 6, were determined for DTF. *QTLA* denotes the locus with the highest LOD score, and largest contributor to DTF variation observed. *QTLB* denotes the second largest contributor to DTF variation. DTF was measured as days to first open flower from emergence. Plants received a 12-h photoperiod of natural daylight. Data are for *n*=167. The 'dotted grey' line denotes the LOD scores derived from *Interval Mapping.* The 'bold black' line denotes the LOD score derived from *MQM* mapping. The 'dotted red' line denotes the chromosome-wide *LOD threshold* of 5.50.

4.3.6.1.1 Effect of DTF1 on time to first open flower in ILL 2601

The individual contributions of each of the two QTL was subsequently analysed by categorising F_2 progeny according to the genotype of their respective peak markers (Figure 4-7). F_2 progeny homozygous for the ILL 2601 allele of the *DTF1* peak marker (DArT-Seq^{TM} 101290509) reported a mean DTF of 37.6 ± 1.08 days. This is 30.0 days earlier ($p < 0.05$) than that observed with progeny homozygous for the ILL 5588 allele of the *DTF1* peak marker (mean DTF = 67.5 ± 4.92) (Figure 4-7B). Progeny heterozygous for the *DTF1* peak marker had an intermediate phenotype (mean DTF = 42.0 ± 1.23), reportedly different ($p < 0.05$) from either homozygous class (Figure 4-7B).

The contribution of *DTF1* was then analysed in only F_2 progeny with an ILL 5588 background for *DTF2.* This was carried out so as to exclude the effect of potential interactions between the early flowering alleles for the peak marker of both *DTF1* and *DTF2*, which can distort the inferred mode of inheritance of *DTF1*.

In this secondary analysis, progeny homozygous for the ILL 2601 allele of the *DTF1* peak marker reported a mean DTF of 41.0 ± 1.81 days, 62.0 days earlier ($p < 0.05$) than progeny homozygous for the ILL 5588 allele of the *DTF1* peak marker (mean DTF = 103 ± 5.91) (Figure 4-7C). Additionally, progeny heterozygous for the *DTF1* peak marker flowered significantly earlier (57.9 days; $p < 0.05$) than progeny homozygous for the ILL 5588 allele of the *DTF1* peak marker (Figure 4-7C). There was no significant difference ($p = 0.539$) observed for DTF between progeny homozygous for the ILL 2601 allele and the heterozygous progeny, suggesting a dominant mode of inheritance for the early-flowering phenotype for *DTF1* in SD (Figure 4-7C), contrary to the earlier finding.

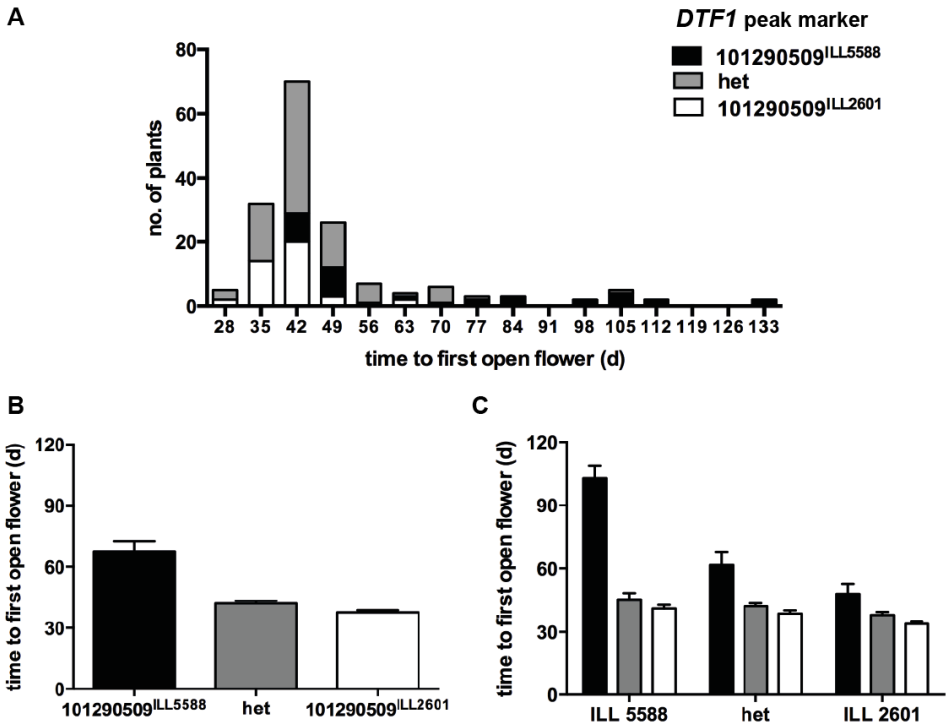


Figure 4-7 Contribution of *DTF1* to early flowering phenology in ILL 2601
(A) DTF segregation of F_2 progeny for the peak *DTF1* marker (DArT-Seq™ marker 101290509) under SD conditions. Data are for *n*=35-90. (B) Association analysis of *DTF1* for DTF under SD conditions. Data are mean ±SE for *n*=35-90. (C) Association analysis of *DTF1* for DTF under SD conditions, against ILL 5588, het, and ILL 2601 background for *DTF2*. Data are mean ±SE for *n*=6-44. DTF was measured as days to first open flower from emergence. Plants received a 12-h photoperiod of natural daylight.

4.3.6.1.2 Effect of DTF2 on time to first open flower in ILL 2601

The contribution of *DTF2* to DTF was subsequently analysed. F_2 progeny homozygous for the ILL 2601 allele of the *DTF2* peak marker (DArT-Seq™ 3659911) at the secondary QTL for flowering time reported a mean DTF of 39.5 ± 11.9 days, 18.1 days earlier ($p < 0.05$) than progeny homozygous for the ILL 5588 allele (mean DTF = 57.5 ± 10.7) (Figure 4-8B). Progeny heterozygous for the *DTF2* peak marker were observed to flower on average within 39.5 ± 11.9 days, with no significance difference ($p = 0.054$) reported between the progeny homozygous for the ILL 2601 allele of the *DTF2* peak marker, suggesting a fully dominant mode of inheritance of the early-flowering phenotype for *DTF2* in SD (Figure 4-8B).

When F_2 progeny with an ILL 5588 background for *DTF1* were only analysed for association, progeny homozygous for the ILL 2601 allele of the *DTF2* peak marker reported a mean DTF of 47.8 ± 4.81 days, 55.2 days earlier ($p < 0.05$) than progeny homozygous for the ILL 5588 allele of the *DTF2* peak marker (mean DTF = 103 ± 5.91) (Figure 4-8C). While, there was a significant difference ($p < 0.05$) between progeny homozygous for the ILL 5588 allele and the heterozygous *DTF2* progeny (mean DTF = 61.62 ± 6.11), no significant difference ($p = 0.0959$) was observed when the heterozygous class was compared to progeny homozygous for the ILL 2601 allele of the *DTF2* peak marker. This is consistent with the earlier observation, and reinforces the dominant mode of inheritance for the ILL 2601 early-flowering phenotype at *DTF2*.

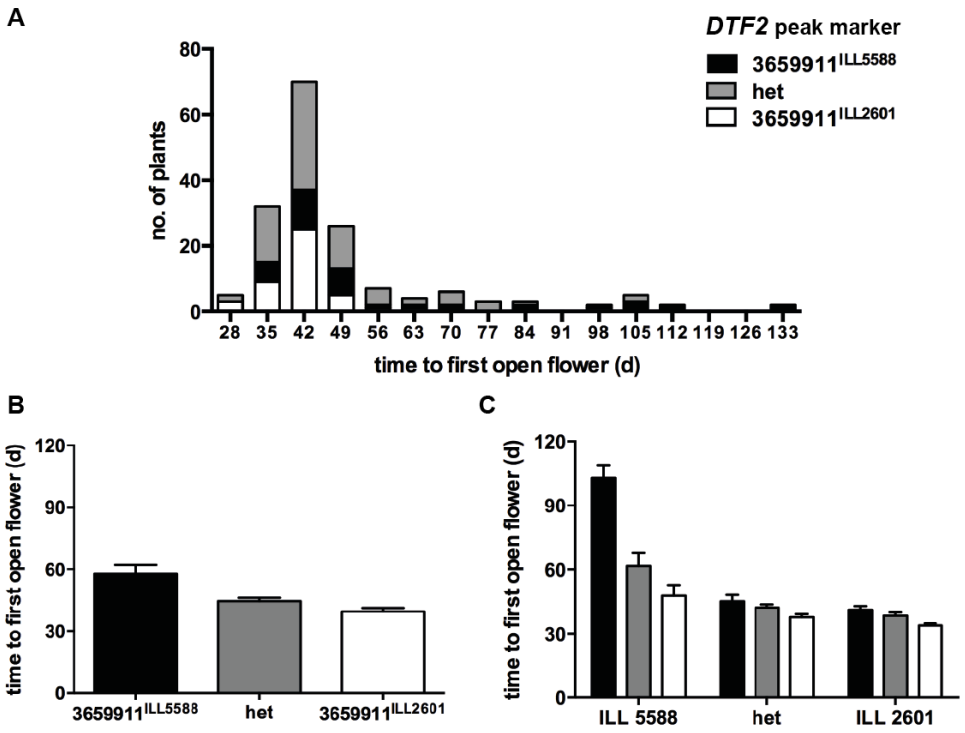


Figure 4-8 Contribution of *DTF2* to early flowering phenology in ILL 2601
(A) DTF segregation of F_2 progeny for the peak *DTF2* marker (DArT-Seq™ marker 3659911) under SD conditions. Data are for *n*=40-82. (B) Association analysis of *DTF2* for DTF under SD conditions. Data are mean ±SE for *n*=40-82. (C) Association analysis of *DTF2* for DTF under SD conditions, against ILL 5588, het, and ILL 2601 background for *DTF1*. Data are mean ±SE for *n*=6-44. DTF was measured as days to first open flower from emergence. Plants received a 12-h photoperiod of natural daylight.

4.3.6.1.3 Interaction between DTF1 and DTF2 for time to first open flower

The interaction between *DTF1* and *DTF2* was examined by comparing the mean DTF for four genotypic combinations for each of the respective peak markers. F_2 progeny heterozygous for either or both loci were excluded from this analysis.

An analysis of variance for the four genotypic classes (Figure 4-9) established that F_2 progeny homozygous for the ILL 5588 allele of both *DTF1* and *DTF2* were observed to flower significantly later ($p < 0.05$) than each of the other three genotypic classes in SD. F_2 progeny homozygous for ILL 2601 alleles from either or both DTF loci were observed to be early flowering. However, there was no significant difference ($p = 0.345$) observed for DTF when comparing progeny homozygous for the ILL 2601 allele at only one DTF loci under SD conditions (Figure 4-9).

The analysis of the individual effects of *DTF1* and *DTF2* additionally suggests an interaction between the two loci for the observed variation in DTF. When in combination, progeny class carrying ILL 2601 alleles at both loci were observed to flower on average, significantly earlier ($p < 0.05$) than progeny homozygous for an ILL 5588 allele at either loci, suggesting that the observed early flowering phenotype is additive. It is therefore likely that both loci act on different pathways to confer ILL 2601 its early phenology, with an extremely early phenotype only attributable to ILL 2601 alleles at both loci (Figure 4-9).

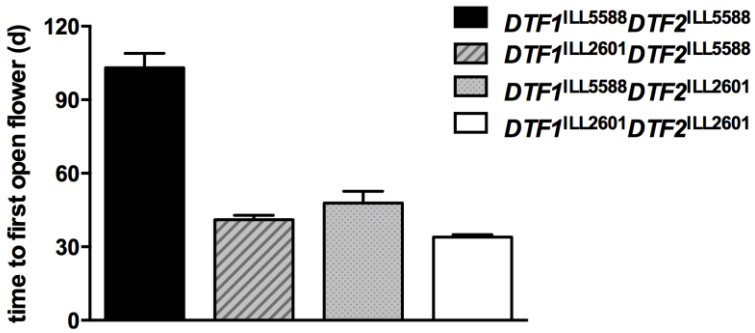


Figure 4-9 Contribution of *DTF1* and *DTF2* to flowering phenology
Four classes of F_2 progeny with homozygous alleles for *DTF1* and *DTF2* (DArT-Seq™ markers 101290509 and 3659911 respectively). Data are mean ±SE for *n*=6-12.

4.3.6.2 Loci contributing to the variation in flowering node

NFD describes the developmental stage of a plant at flowering. In this study, a strong positive correlation (R^2_{adj} = 0.905) between both DTF and NFD is determined (Figure 4-3). Additionally, similar to observations in ILL 5588 under SD described in 4.3.1, it was observed that in some F_2 progeny the NFD was not always the node at which the first floral structure was initiated. Instead in these plants the nodes bearing aborted or undeveloped flowers preceded the appearance of a fully developed open flower (Figure 4-1A).

A preliminary QTL analysis of NFD data determined three loci, two on linkage group 6, and a third on linkage group 2, for the observed variation in NFD (Table 4-3 and Figure 4-11). For the purpose of this thesis, the three NFD loci have been provisionally assigned *NFD1*, *NFD2* and *NFD3*, in the order of their contribution to the observed variation. Both *NFD1* and *NFD2* were determined to be co-located with *DTF1* and *DTF2* respectively (Table 4-3), suggesting that it is likely a single gene at each of these loci is responsible for both quantitative traits. No DTF locus was identified at *NFD3*.

A secondary QTL analysis was subsequently performed using NFI data to identify QTL contributing to the trait, and determine the relationship between NFD and NFI. One locus was identified to contribute to the observed variation in NFI in the ILL 2601 x ILL 5588 F_2 population. For the purpose of this thesis, the NFI locus has been provisionally assigned *NFI1*. *NFI1* has a maximum LOD score of 8.00, and is estimated to contribute to 20.0% of the observed variation for NFI in the F_2 population (Table 4-3). *NFI1* was also determined to be co-located with both *DTF1* and *NFD1* on linkage group 6 NFD (Table 4-3 and Figure 4-11).

A third round of QTL analysis was performed using DFD data to identify QTL contributing to the trait, and determine the relationship between NFD and DFD. Two loci were identified to contribute to the observed variation in DFD in the ILL 2601 x ILL 5588 F_2 population. For the purpose of this thesis, the DFD loci have been provisionally assigned *DFD1* and *DFD2*, in the order of their contribution to

the observed variation. *DFD1* has a maximum LOD score of 11.0, and is estimated to contribute to 14.6% of variation for DFD in the segregating population (Table 4-3). *DFD2* has a reported maximum LOD score of 6.00, and is estimated to contribute to 7.40% of variation for DFD in the segregating population (Table 4-3). *DFD1* was co-located (±1 cM) with *DTF2* and *NFD2* on linkage group 6, and *DFD2* was co-located (±1 cM) with *NFD3* on linkage group 2 (Table 4-3 and Figure 4-11).

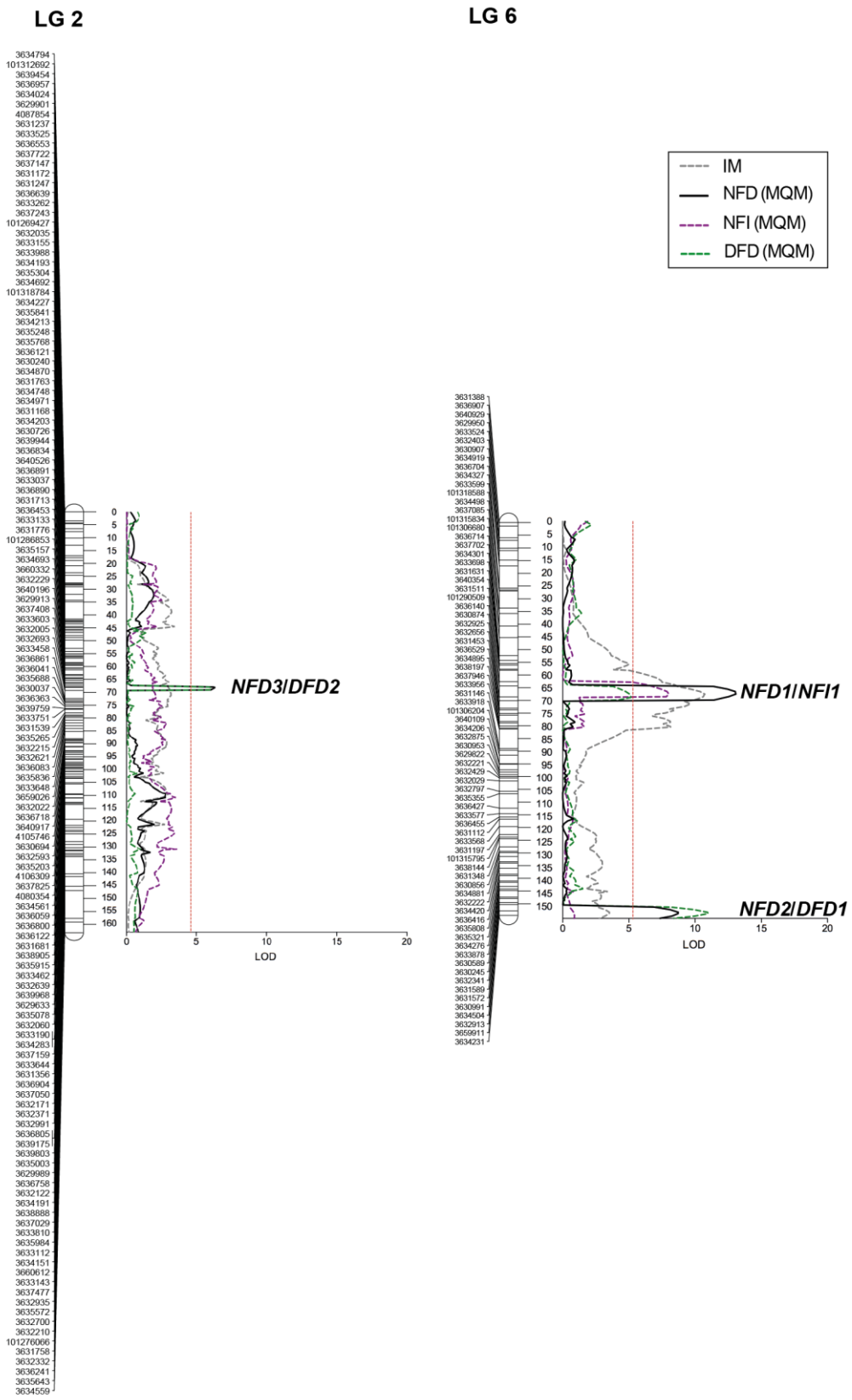


Figure 4-10 Flowering node loci in ILL 2601 x ILL 5588 F_2 population

One locus on lentil linkage group 2 and two loci on lentil linkage group 6 were determined in the QTL analysis for NFD evaluated under SD. The 'dotted grey' line denotes LOD scores derived from *Interval Mapping* for NFD. The 'bold black' line denotes LOD scores derived from *MQM* mapping for NFD. The 'dotted purple' line denotes LOD scores derived from *MQM* mapping for

NFI. The 'dotted green' line denotes LOD scores derived from *MQM* mapping for DFD node interval. The 'dotted red' line denotes the chromosome-wide *LOD threshold* of 4.60 for NFD in LG 2 and 5.30 for NFD in LG 6. The genome-wide *LOD threshold* for NFD was determined at 5.30. The chromosome-wide *LOD threshold* for NFI was determined at 5.10. The chromosome-wide *LOD threshold* for DFD was determined at 4.60 and 5.30 for linkage group 2 and 6 respectively. Plants received a 12-h photoperiod of natural daylight. Data are for *n*=165.

4.3.6.2.1 Effect of NFI1 on node of floral initiation

No locus has been previously reported nor described for its role in controlling NFI in lentil. As a trait, NFI is proposed in 4.3.1 to not be responsive to photoperiod (Figure 4-1E), and is proposed to be controlled by a single locus *NFI1* in the ILL 2601 x ILL 5588 F_2 population (Table 4-3 and Figure 4-11). *NFI1* is additionally suggested to be co-located with both *NFD1* and *DTF1* (Table 4-3). To further inform the role of this locus in conferring earliness to ILL 2601, an association analysis for the trait was carried out.

The effect of *NFI1* on NFI was analysed by categorising F_2 progeny according to the genotype of the peak marker. F_2 progeny homozygous for the ILL 2601 allele of the *NFI1* peak marker (DArT-Seq^{TM} 3631511) reported a mean NFI of 15.6 ± 0.239 nodes, 3.01 nodes earlier ($p < 0.05$) than progeny homozygous for the ILL 5588 allele (mean NFI = 18.64 ± 0.803 nodes) (Figure 4-11A). While there was a significant difference in NFI observed between the heterozygous class (mean NFI = 15.85 ± 0.169 nodes) and the progeny homozygous for the ILL 5588 allele, no significant difference ($p = 0.387$) was observed between the heterozygous class and progeny homozygous for the ILL 2601 allele for NFI.

The NFI means reported for each genotypic class suggests, that the ILL 2601 *NFI1* allele confers F_2 progeny with a lower node for the initiation of the first floral structure (Figure 4-11A). Furthermore, the inheritance of the early phenotype is observed to be dominant (Figure 4-11A). This observation is consistent with the dominant nature of the early flowering DTF phenotype conferred by the ILL 2601 allele of *DTF1* (Figure 4-7B).

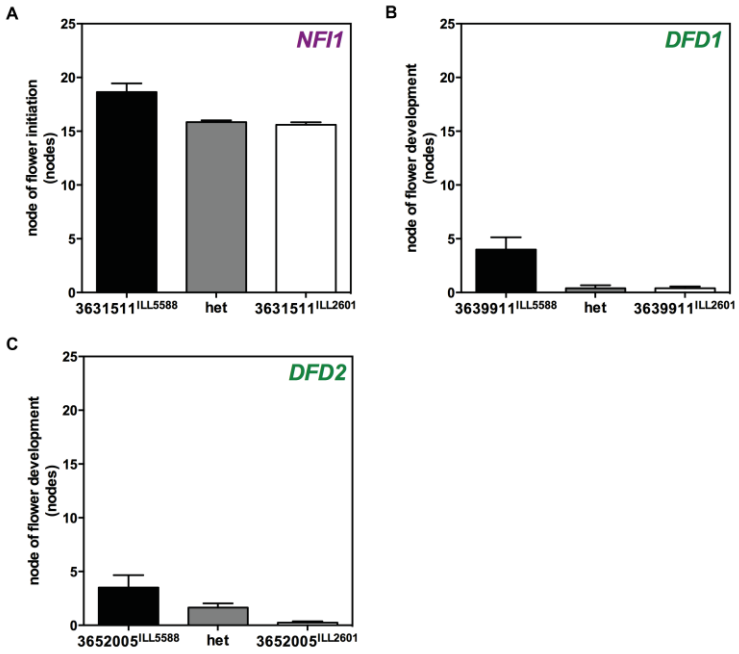


Figure 4-11 Characterisation of the node of first open flower
(A) Effect of *NFI1* peak marker DaRT-Seq™ 3631511 on NFI under SD conditions. Data are for *n*=36-87. (B) Effect of *DFD1* peak marker DaRT-Seq™ 3639911 on DFD under SD conditions. Data are mean ±SE for *n*=39-82. (C) Effect of *DFD2* peak marker DaRT-Seq™ 3632005 on DFD under SD conditions. Data are mean ±SE for *n*=36-88. Plants received a 12-h photoperiod of natural daylight.

4.3.6.2.2 Effect of DFD1 and DFD2 on node of flower development

Similar to NFI, no locus has also been previously reported nor described for its role in the regulation and control of DFD in lentil. Two loci were determined to contribute to the DFD variation in the ILL 2601 x ILL 5588 F_2 population. To further elucidate the role of these loci in conferring earliness to ILL 2601, an association analysis for the trait was carried out.

The effect of each of the two loci regulating DFD was analysed by categorising F_2 progeny according to the genotype of their respective peak markers. In the analysis of F_2 progeny homozygous for the ILL 2601 allele of the *DFD1* peak marker (DArT-Seq™ 3639911), a mean interval of 0.39 ± 0.166 nodes, significantly shorter ($p < 0.05$) than progeny homozygous for the ILL 5588 allele (mean DFD = 4.00 ± 1.14 nodes) was reported (Figure 4-11B). There was no significant difference ($p = 0.0753$) reported between the heterozygous class (mean DFD = 1.12 ± 0.286 nodes) and progeny homozygous for the ILL 2601 allele of the *DFD1* peak marker, suggesting that the shorter node interval

between floral induction and first open flower, characteristic of the early-flowering ILL 2601 in SD is dominantly inherited at *DFD1* (Figure 4-11B). This is consistent with the observed mode of inheritance of the early DTF phenotype conferred by *DTF2* in SD.

To exclude the effect of potential interactions between the DFD loci, F_2 progeny with an ILL 5588 background for the secondary locus *DFD2* were only analysed for association to observed delay to developed flowers. It was observed that progeny homozygous for the ILL 2601 allele of the *DFD1* peak marker reported to a mean DFD of 0.750 ± 0.413 nodes, significantly shorter ($p < 0.05$) than progeny homozygous for the ILL 5588 allele of the *DFD1* peak marker (mean DFD = 15.8 ± 4.36 nodes). While, there was a significant difference ($p < 0.05$) between progeny homozygous for the ILL 5588 allele and the heterozygous class (mean DFD = 1.27 ± 0.597 nodes), no significant difference ($p = 0.478$) was observed when the heterozygous class was compared to progeny homozygous for the ILL 2601 allele. This is consistent with earlier observations and the observed mode of inheritance for *DFD1* in SD.

The contribution of *DFD2* to the observed variation for DFD was then analysed. F_2 progeny homozygous for the ILL 2601 allele of the *DFD2* peak marker reported a mean DFD interval of 0.27 ± 0.126 nodes, significantly shorter ($p < 0.05$) than progeny homozygous for the ILL 5588 allele (mean DFD = 3.06 ± 1.07 nodes) (Figure 4-11C). While there was a significant difference ($p < 0.05$) observed between the heterozygous class (mean DFD = 1.64 ± 0.397 nodes) and homozygous progeny, no significant difference ($p = 0.126$) was observed between the heterozygous class and progeny homozygous for the ILL 5588 allele (Figure 4-11C). This suggests that the extended DFD node interval conferred by the ILL 5588 allele for *DFD2* is dominant in SD. This is in contrast to the mode of inheritance for DFD conferred by *DFD1*.

However, when F_2 progeny with an ILL 5588 background for *DFD1* were only analysed for association to DFD, the heterozygous class (mean DFD = 4.06 ± 1.42

nodes) was observed to be an intermediate between both homozygous classes, with a significant difference ($p < 0.05$) observed for DFD for both heterozygous classes. Progeny homozygous for the ILL 2601 allele of *DFD2* reported a mean DFD of 15.8 ± 4.36 nodes. Progeny homozygous for the ILL 5588 allele reported a mean DFD of 0.250 ± 0.112 nodes. Contrary to the earlier analysis, in this analysis the inheritance of the observed delay to developed flowers, as conferred by *DFD2* appears to be co-dominant when the effect of ILL 2601 alleles at *DFD1* are excluded.

4.3.6.2.3 Interaction between QTLB and QTLC for DFD

An analysis of variance for the four classes (Figure 4-12) reported that a significant difference ($p < 0.05$) in DFD exists between progeny homozygous for the ILL 5588 allele at both *DFD1* and *DFD2*, and each of the three other classes. There was however, no significant difference observed for DFD between the classes carrying ILL 2601 alleles for either or both loci. This suggests that both *DFD1* and *DFD2* are complementary to each other, and that the early-flowering habit conferred by the ILL 2601 alleles is a consequence of genes functioning within the same induction pathway.

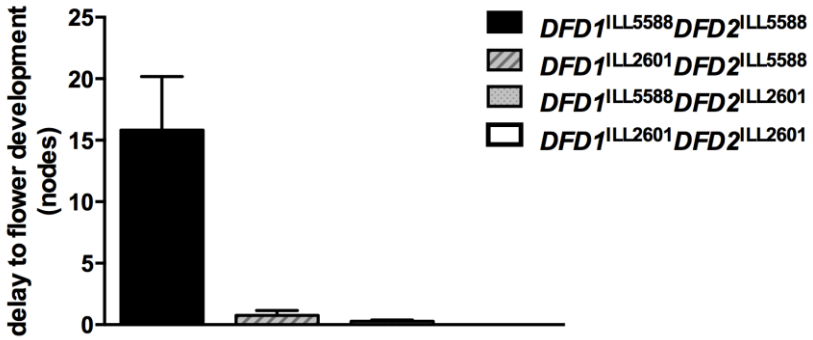


Figure 4-12 Contribution of *DFD1* and *DFD2* to flowering phenology
Four classes of F_2 progeny with homozygous alleles for *DFD1* and *DFD2* (DArT-Seq™ markers 3659911 and 3632005 respectively). Data are mean ±SE for *n*=5-16.

4.3.6.3 Loci contributing to the variation in emergence time

A bimodal segregation for DTE was observed (Figure 4-13A) for the F_2 progeny, with an early class emerging between 11 and 15 days after sowing, followed by a later class with a wider emergence range from 15 to 31 days after sowing.

Additionally, no linear relationship between DTE and DTF was determined (R^2_{adj} = 0.009), suggesting that both traits are independent of each other. To investigate the genetic basis for this variation, a QTL analysis was undertaken.

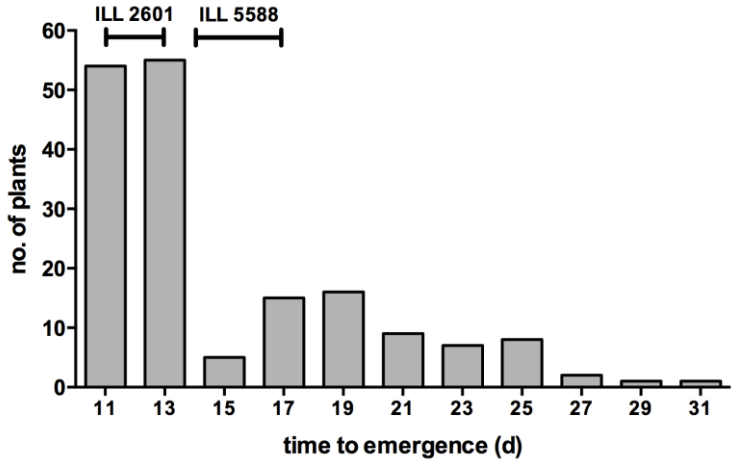


Figure 4-13 Days to emergence in ILL 2601 x ILL 5588 F_2 population
Distribution of F_2 progeny for DTE. DTE for ILL 2601 and ILL 5588 are indicated on histogram. Data are for *n*=173.

Two loci were identified for DTE during QTL analysis, one on linkage group 7 and another on linkage group 5 (Figure 4-14). For the purpose of this chapter, these loci will be provisionally referred to as *DTE1* and *DTE2* respectively.

DTE1 reported a maximum LOD score of 34.5, is estimated to account for 52.8% of observed variation for DTE. The second loci *DTE2* reported a maximum LOD score of 11.5, and is estimated to account for 12.6% of observed variation for DTE. In both instances, the ILL2601 allele for the peak marker was associated with reduced DTE.

A chromosome-wide *LOD threshold* of 5.20 and 4.90 was employed for the determination of the QTL in linkage group 5 and 7 respectively. The genome-wide *LOD threshold* reported was 7.50.

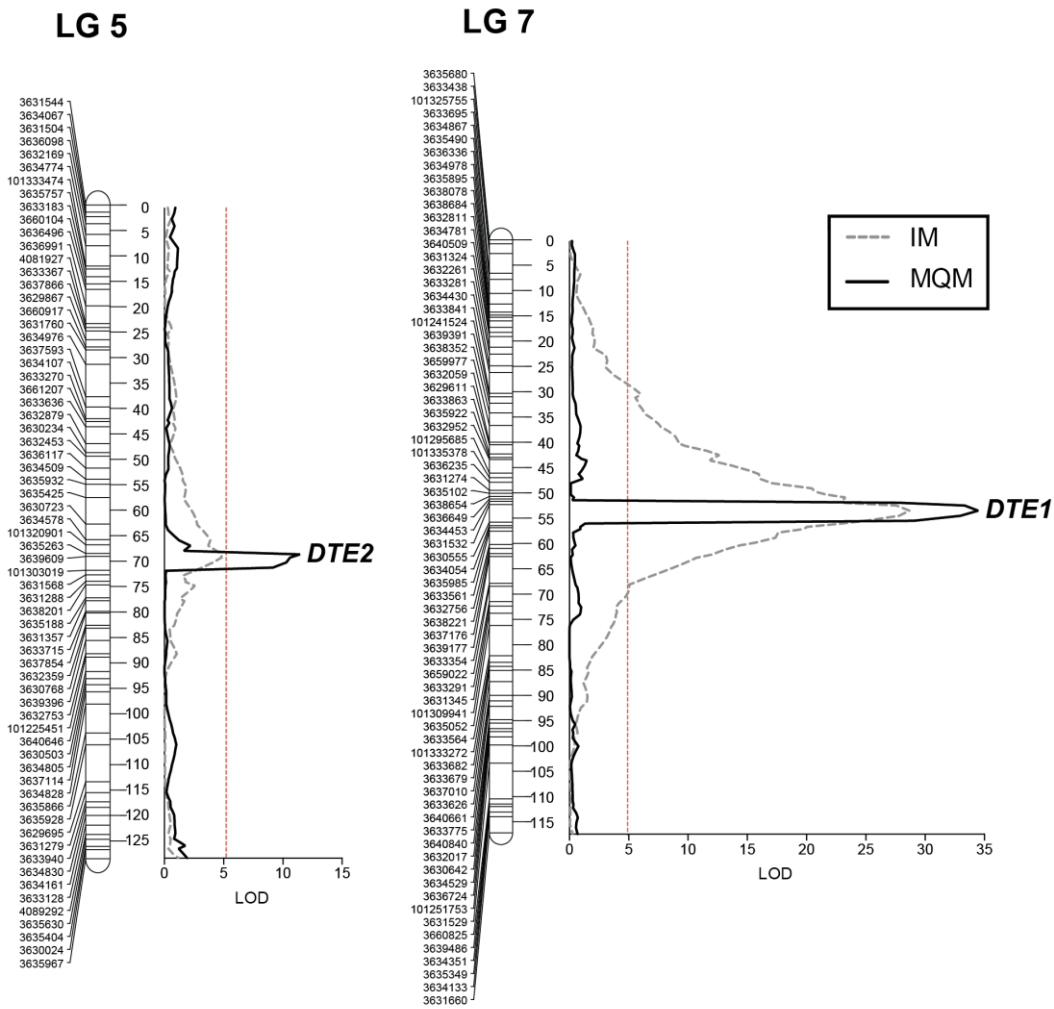


Figure 4-14 Emergence time loci in ILL 2601 x ILL 5588 F_2 population
Two loci, located on linkage group 5 and linkage group 7, were identified in the QTL analysis for DTE evaluated under SD. *DTE1* denotes the locus with the highest LOD score, and largest contributor to DTE variation observed. *DTE2* denotes the second locus contributing to observed DTE variation. DTE measured as the interval between sowing and the appearance of the first pair of fully expanded bifoliate leaves. Plants received a 12-h photoperiod of natural daylight. Data are for *n*=173. The 'dotted grey' line denotes the LOD scores derived from *Interval Mapping.* The 'bold black' line denotes the LOD score derived from *MQM* mapping. The 'dotted red' line denotes the chromosome-wide *LOD threshold* of 5.20 and 4.90 for linkage groups 5 and 7 respectively. The genome-wide wide *LOD threshold* was 7.50.

4.3.6.3.1 Effect of DTE1 on emergence time

The effect of *DTE1* on the observed variation for DTE was analysed by categorising F_2 progeny according to the genotype of their respective peak markers (Figure 4-15). F_2 progeny homozygous for the ILL 2601 allele (mean DTE = 12.0 ± 0.287 days) of the *DTE1* peak marker (DArT-Seq™ 3631532) emerged 9.2 days earlier ($p < 0.05$) than progeny carrying the ILL 5588 allele (mean DTE = 21.2 ± 0.745 days) (Figure 4-15B). Heterozygous progeny (mean DTE = 13.6 ± 0.338 days) were analysed to be an intermediate ($p < 0.05$) between that of the two homozygous groups, suggesting incomplete dominance (Figure 4-15B).

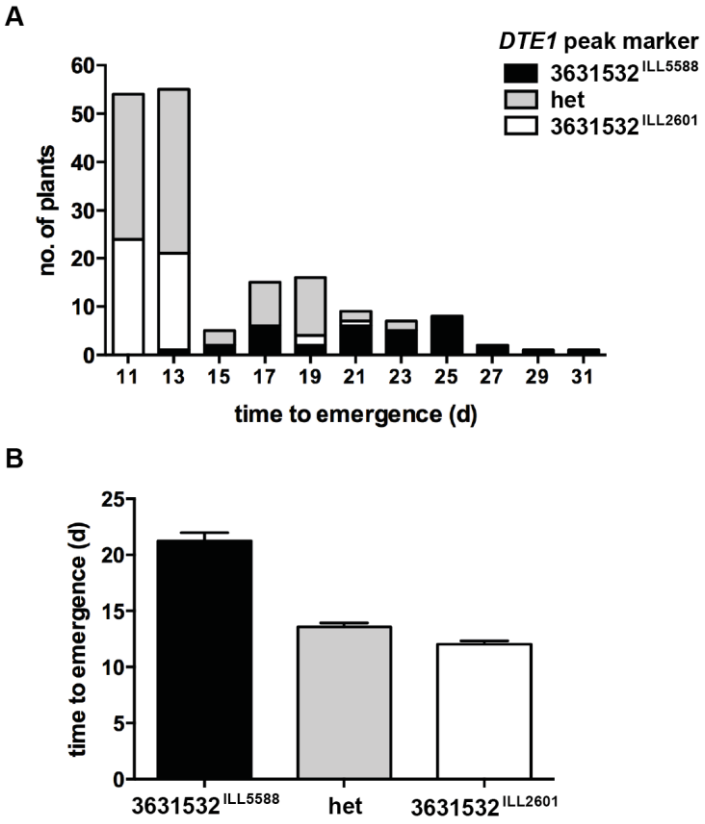


Figure 4-15 Contribution of *DET1* to truncated pre-emergent phase in ILL 2601
(A) DTE segregation of F_2 progeny for the peak *DTE1* marker (DArT-Seq™ marker 3631532) under SD conditions. Data are for *n*=173. (B) Association analysis of *DTE1* for DTE under SD conditions. Data are mean ±SE for *n*=34-92. DTE was measured as the period interval between sowing and the appearance of the first pair of fully expanded bifoliate leaves. Plants received a 12-h photoperiod of natural daylight.

4.3.6.3.2 Effect of DTE2 on emergence time

The effect of *DTE2* on the observed variation for DTE was subsequently analysed by categorising F_2 progeny according to the genotype of their respective peak markers (Figure 4-17). F_2 progeny homozygous for the ILL 2601 allele (mean DTE = 13.6 ± 0.611 days) of the *DTE2* peak marker (DArT-Seq™ 3635263) emerged 4.12 days earlier ($p < 0.05$) than progeny carrying the ILL 5588 allele (mean DTE = 17.7 ± 0.851 days) (Figure 4-17B). Additionally, heterozygous progeny (mean DTE = 14.0 ± 0.430 days) did not emerge significantly later ($p = 0.603$) earlier than progeny homozygous for the ILL 2601 allele, suggesting a dominant mode of inheritance for the ILL 2601 allele at *DTE2* (Figure 4-17B).

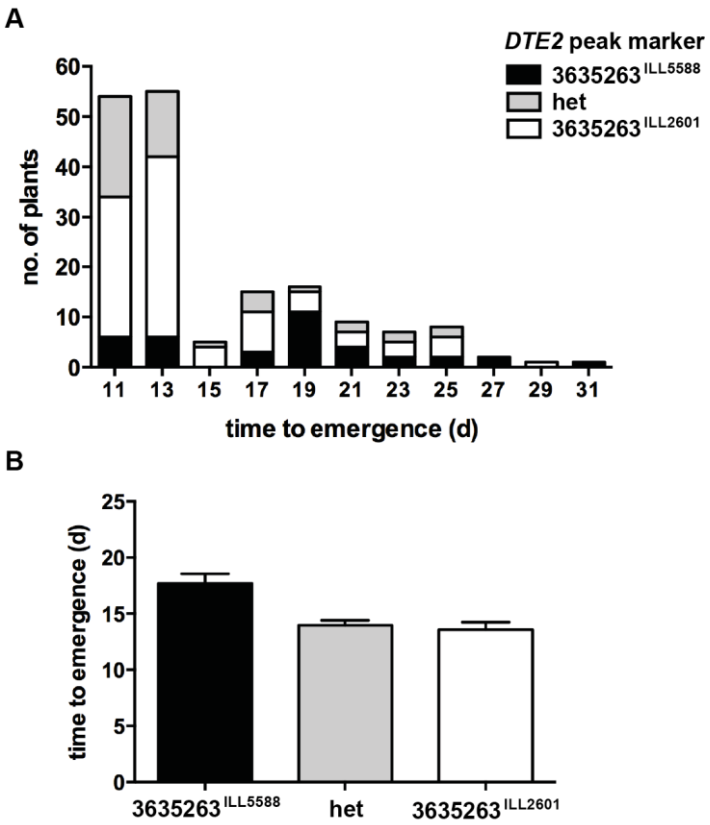


Figure 4-16 Contribution of *DET2* to truncated pre-emergent phase in ILL 2601
(A) DTE segregation of F_2 progeny for the peak *DTE2* marker (DArT-Seq™ marker 3635263) under SD conditions. Data are for *n*=173. (B) Association analysis of *DTE1* for DTE under SD conditions. Data are mean ±SE for *n*=37-91. DTE was measured as the period interval between sowing and the appearance of the first pair of fully expanded bifoliate leaves. Plants received a 12-h photoperiod of natural daylight.

4.3.6.3.3 Interaction between DTE1 and DTE2 for emergence time

An analysis of variance between the four classes revealed a significant difference ($p < 0.05$) in DTE between progeny homozygous for the ILL 5588 allele at both *DTE1* and *DTE2*, and each of the three other classes with homozygous ILL 2601 alleles at either or both loci conferring a truncated pre-emergent phase. Moreover, it was observed that progeny carrying the ILL 2601 alleles at *DTE1* had a shorter DTE when compared to progeny carrying the ILL 2601 alleles at *DTE2*. However when in combination, the effect of ILL 2601 alleles at both DTE loci was not additive. Instead, no significant difference ($p = 0.080$) was observed between progeny homozygous for the ILL 2601 allele at *DTE2* and progeny homozygous for the ILL 2601 allele at both loci. This suggests that both *DTE1* and *DTE2 are* epistatic, and that it is likely that both function in the same pathway.

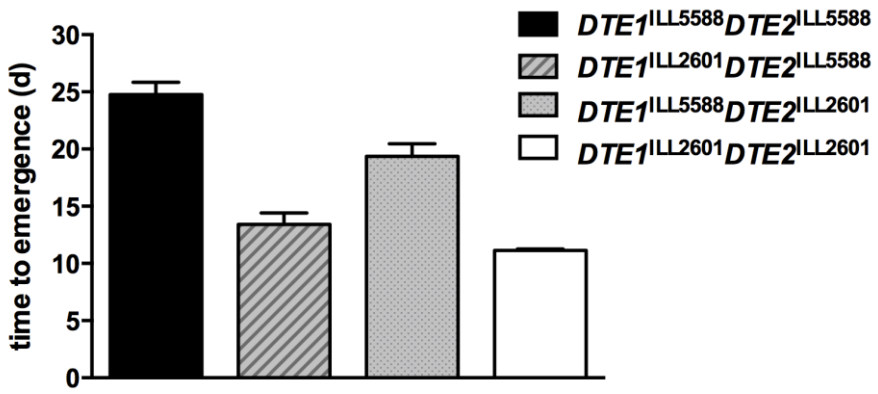


Figure 4-17 Interaction between DTE1 and DTE2 for DTE
Mean time to emergence for the four possible combinations of homozygous *DTE1* and *DTE2* genotypes, as inferred from peak markers 3631532 and 3635263, respectively. Data are mean ±SE for *n*=8-12.

4.3.6.4 QTL co-location for early traits

The co-location of loci associated with traits conferring earliness including DTF, NFD, NFI, and DFD, was observed in this study (Table 4-3). In all reported instances, the locus was observed to either occur at the same position of the chromosome or within ±1 cM of the co-located locus.

It is proposed, from this chapter, that it is likely that *DTF1*, *NFD1*, and *NFI1* are controlled by the same locus. For the purpose of this thesis, this locus will be assigned *QTLA*. It is also proposed that *DTF2*, *NFD2*, and *DFD1* are similarly controlled by a single locus. For the purpose of this thesis, this locus will be assigned *QTLB*. The third locus *DFD2*, attributable to the variation in the delay to a developed flower, will be assigned *QTLC* for consistency. *QTLC* had no detectable contribution to the observed variation for DTF.

4.3.7 Mapping of other quantitative traits

QTL for several other quantitative traits were also identified in the ILL 2610 x ILL 5588 F_2 population (Table 4-4). Traits mapped included plant height (PH), internode interval between nodes 1 and 9 (IN9), and number of early branches (EBN) and total length of early branches (EBL) (Figure 4-18). None of these traits were observed to occur within the QTL confidence intervals for the loci controlling earliness *QTLA* and *QTLB*, or at *QTLC*. It was however observed all three described quantitative traits were controlled partially by the *DTE1* locus, defined to be involved in the control of the pre-emergent phase.

	QTL information						Peak Marker Information		
Trait	**QTL**	**LG**	**Position (cM)**	**LOD threshold**	**Max LOD**	**% Expl. Variation**	**DArT-Seq™ Marker ID**	**Marker position (cM)**	**Marker LOD**
Number of early branches (EBN)	*EBN1*	7	53.3	4.4	8.5	13.5	3631532	52.3	8.3
	EBN2	5	77.2	5.1	5.3	8.1	3635188	77.2	5.3
Total length of early branches (EBL)	*EBL1*	2	47.8	6.0	6.5	14.3	3630240	46.8	6.5
	EBL2	7	52.3	5.0	5.2	11.1	3631532	52.3	5.2
Internode interval between nodes 1 and 9 (IN9)	*IN91*	7	53.3	4.2	11.0	22.1	3631532	52.3	10.9
	IN92	7	100.8	4.2	5.1	9.5	101251753	99.8	5.1
Plant height (PH)	*PH1*	7	54.3	4.7	13.0	29.2	3630555	55.8	12.5

Table 4-4 Other quantitative traits mapped in ILL 2601 x ILL 5588 F_2 population
Number of early branches (EBN), total length of early branches (EBL), internode interval between nodes 1 and 9 (IN9), and plant height (PH) mapped. Max LOD refers to maximum LOD score for each trait determined using MQM. % Expl. Variation refers to percentage of total variation for a particular trait attributable to a QTL, determined using MQM. Marker LOD refers to LOD score for peak markers for specific traits determined using MQM.

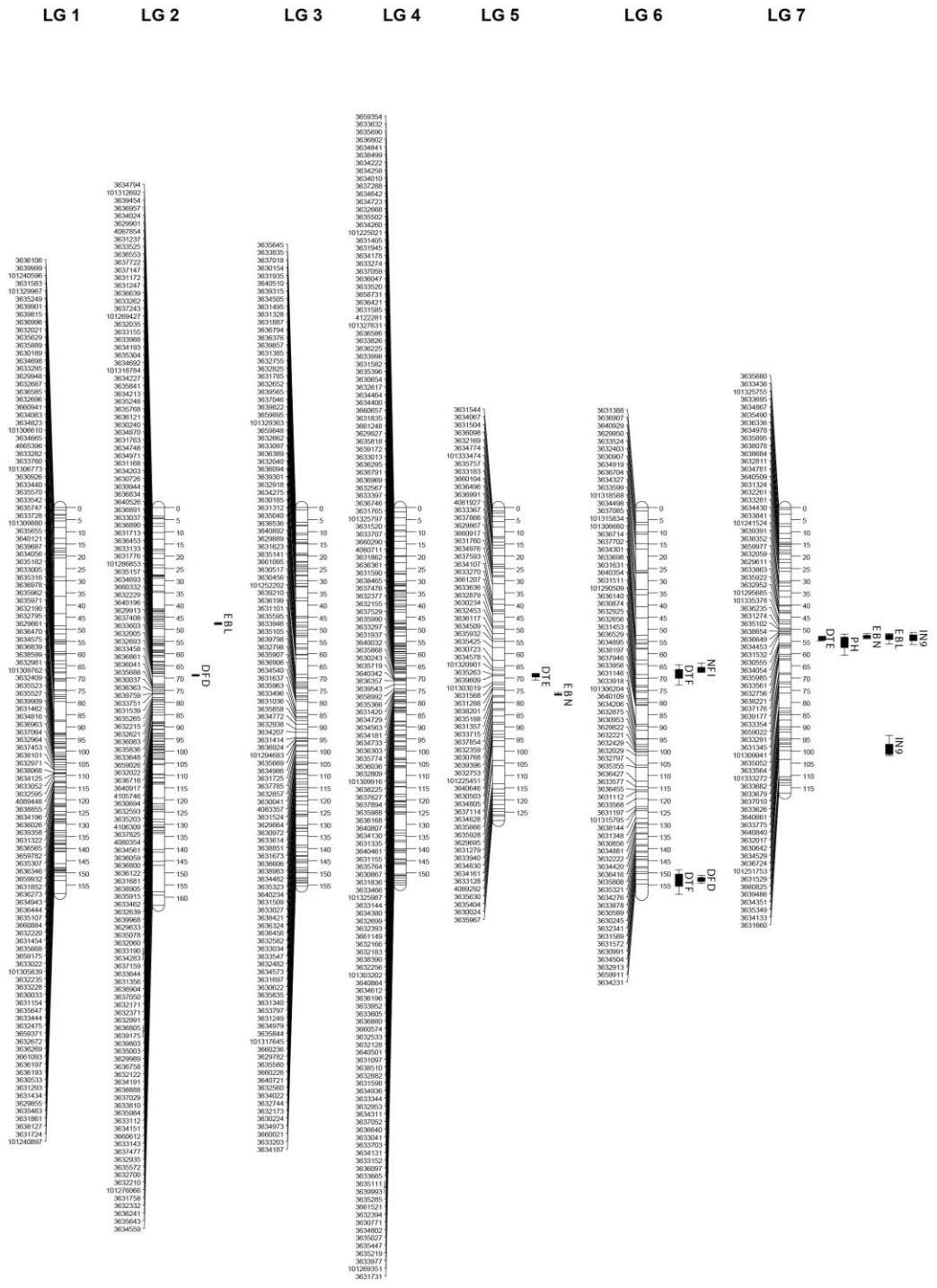


Figure 4-18 ILL 5588 x ILL 2601 F_2 genetic linkage map

Genetic linkage map consists of seven contiguous linkage groups corresponding to the seven chromosomes of the *Lens* genus. 1-LOD and 2-LOD intervals are marked for each quantitative trait; days to first open flower (DTF), node of first floral structure (NFI), interval between node of first floral structure and node of flower development (DFD), plant height (PH), total branches at three weeks from emergence (EBN), total length of branches at three weeks from emergence (EBL), internode length between nodes 1 and 9 (IN9).

4.4 Discussion

This chapter sought to determine the genetic basis for the early habit of an Indian landrace. ILL 2601 is amongst the earliest flowering in the lentil germplasm, and is likely to represent the fullest extent of the adaptive early flowering phenotype of the *pilosae* lentil (Erskine, W. et al., pers. comm; Weller and Murfet, unpublished). In this chapter it was determined that ILL 2601 is not responsive to photoperiod and carries the functional lentil *Sn*/*ELF3*. To understand the genetic basis for this phenotype, a F_2 population segregating for flowering time was established, a genetic linkage map constructed, and loci contributing earliness identified through QTL analyses.

4.4.1 Genetic control of earliness in ILL 2601

The phenotypic characterisation of ILL 2601 dissected the early habit of the Indian landrace into three major quantitative traits, namely *time to emergence* (DTE) from sowing, *days to flowering* (DTF) from emergence, and the *node of flower development* (NFD). The latter was further dissected into two independent traits, *node of floral initiation* (NFI) and *delay to flower development* (DFD). The former is not responsive to prevailing photoperiod, while the later was only observed to occur under non-inductive long days in ILL 5588 (photoperiod-sensitive accession).

To probe the genetic basis for these traits, this chapter established a F_2 population segregating for flowering time with ILL 5588. The segregants were genotyped and a genetic linkage map was constructed using DArT-Seq™ markers. Through QTL mapping it was determined that the earliness observed in ILL 2601 relative to ILL 5588 is a function of at least five different loci. Two loci were identified to contribute to the variation for the pre-emergent phase (DTE), and three loci, namely *QTLA*, *QTLB*, and *QTLC* (refer to 4.3.6.4), were identified to collectively contribute to the variation for time (DTF) and node (NFD) for the transition to flowering. It was also determined that DTE and DTF are independent of each other.

4.4.1.1 Genetic control of the pre-emergent phase

The period of the pre-emergent phase, designated *days to emergence* (DTE) in this study, has not been reported in lentil to contribute to an early phenotype. The pre-emergent phase in lentil is described as the period between sowing and emergence (Roberts et al., 1986). QTL analysis for this trait determined two loci, namely *DTE1* and *DTE2*, responsible for the observed variation for DTE. Moreover, it was determined that the loci are complementary and are likely to act on the same pathway. Furthermore, ILL 2601 alleles for both loci confer a shift to an early phenotype. Interestingly, no DTF or flowering node (NFD, NFI, and DFD) loci were determined to be co-located with either DTE locus.

Associated with germination time and seed dormancy, the genetic control of this trait is suggested to be regulated by a single dominant gene controlling the hard seed coat (Ladizinsky, 1985). Ladizinsky (1985) adds that this trait can be overcome by seed coat scarification. Roberts et al. (1986) has also suggested that in lentil this phase is controlled by the germination rate, which is determined to be a function of temperature (Covell et al., 1986).

However, in this study, the variation for DTE cannot be attributed to the seed coat as the seed coat tissue is of maternal origin and hence genetically F_1. Furthermore, as described in Section 2.1, all seed coats were scarified and seeds imbibed prior to sowing. This pre-sowing seed treatment further excludes the role of the seed coat in the observed variation for DTE.

Apart from work relating to the hard seed coat and its role in regulating the pre-emergent phase, there is no precedence for genetic work on germination time in lentil. In *M. truncatula*, one loci located on chromosome 8, (corresponding to lentil linkage group 7), and two loci on chromosome 5, (corresponding to lentil linkage group 5), have been previously implicated in the control of germination time, and the pre-emergent growth phase (Dias et al., 2011). Work in *M. truncatula* affords basis for future work in lentil relating to the molecular resolution of these DTE loci.

4.4.1.2 Genetic control of flowering time and flowering node

The genetic basis for the early-flowering phenotype of the *pilosae* lentil is not known. It was determined in this study that the variation for flowering time and node in the ILL 2601 x ILL 5588 F_2 population is controlled by multiple major loci, with ILL 2601 alleles at *QTLA*, *QTLB*, and *QTLC* affording major shifts to an early flowering phenotype.

Observations of the photoperiodic response of ILL 2601 and ILL 5588, and QTL mapping for DTF, NFD, NFI, and DFD in this study point to a photoperiod-independent and a photoperiod-dependent basis for the control of the flowering phenotype. This implies that it is likely that multiple pathways for flowering occur in lentil, consistent with observations in other legume systems (Weller and Ortega-Martinez, 2015). This study also proposes that the altered regulation of each of these pathways by one or more loci can synergistically afford a shift in the flowering phenotype.

In this study, it was determined that the photoperiod-independent regulation of the flowering phenotype in lentil is controlled by a single locus at *QTLA*. *QTLA* functions to regulate both the interval (time) between DTE and NFD, and the developmental node for NFI, while affording ILL 2601 a dominantly inherited early-flowering phenotype. It is not known how *QTLA* is regulated, or if polymorphisms in the ILL 2601 allele for *QTLA* result in a loss-of-function or a gain-of-function mutation, or if the locus is regulated by specific environmental stimuli. A photoperiod-independent, dominantly inherited early-flowering phenotype has been described by Jaudal et al. (2013) for *M. truncatula*. This will be explored in Chapter 5.

This chapter also identified that in the studied population, the photoperiod-dependent regulation of the flowering phenotype is a function of two epistatic loci, namely *QTLB* and *QTLC*. *QTLB* appears to regulate both the interval (time) between DTE and a fully developed flower, and the interval (node) between NFI and NFD (DFD). *QTLC* conversely is determined to only contribute to the

variation for DFD. *QTLB* and *QTLC* are complementary to each other, and ILL 2601 alleles at either locus confer progeny an early flowering phenotype. Both loci complement the photoperiod-independent *QTLA* to synergistically shift the flowering phenotype. Chapter 3 determined that that the lentil *Sn* functions to confer photoperiod-sensitivity, and is an *Arabidopsis ELF3* orthologue. In this chapter, the mutant *elf3-1* was determined to not contribute to the photoperiod-insensitivity of ILL 2601. Chapter 5 will further explore the molecular basis for *QTLB* and *QTLC*.

4.4.2 Genetic linkage map, macrosynteny with Medicago, and coverage

A high-density genetic linkage map with seven linkage groups corresponding to the seven chromosomes of the lentil genome was constructed in this study. Prior to this study, only one other gene-based intraspecific linkage map for lentil (Sharpe et al., 2013), with seven linkage groups has been reported.

The genetic linkage map constructed for the ILL 2601 x ILL 5588 F_2 population formed the basis for QTL mapping for quantitative traits evaluated in this study. It was therefore imperative that the linkage map constructed had good coverage, with few large intervals between markers. The lentil genetic linkage map reported in this chapter had an average density of 1.41 markers per cM, with only one pair of markers reporting an interval exceeding 10 cM. This compares favourably to work undertaken by Sharpe et al. (2013) where a density of 1.06 markers per cM and several large intervals exceeding 10 cM was reported.

The ILL 2601 x ILL 5588 genetic linkage map developed in this study has also further informed the macrosyntenic relationship between lentil and *M. truncatula* and further developed our understanding of the lentil genome. The genetic linkage map presented describes major inversions for regions of lentil linkage groups 1 and 7, when compared to corresponding regions in *M. truncatula*. This is consistent with findings by Sharpe et al. (2013).

In Sharpe et al. (2013) the translocation of *M. truncatula* chromosome 6 to lentil linkage group 2 was described. In the presented linkage map a large translocation of *M. truncatula* chromosome 6 to lentil linkage group 2 is noted, although only a few markers with sequence similarity to *M*. *truncatula* orthologues in chromosome 6 were determined. More work is required to resolve the extent of translocation and the order of genes in this region of lentil linkage group 2.

The genetic linkage map also presented translocations between the ends of the lentil linkage groups 4 and 8, when compared to corresponding regions in *M*. *truncatula*. The observed difference is due to an aberrant chromosomal arrangement that has resulted from reciprocal translocations of the long arms of chromosomes 4 and 8 in the *M. truncatula* model accession A17 (Kamphuis et al., 2007). Interestingly, Sharpe et al. (2013) did not report similar translocations for lentil. The ILL 2601 x ILL 5588 genetic linkage map also proposes that the lentil linkage group 3 is collinear with *M*. *truncatula* chromosome 3, contrary to the findings by Sharpe et al. (2013) which described a large inversion.

It is not clear from the literature which genome release version of *M*. *truncatula* was utilised in Sharpe et al. (2013). It is likely that their use of a different version to that used in this study is the reason for the discrepancies in reported synteny.

4.4.3 Next steps

This chapter explored the genetic basis for the observed earliness in ILL 2601. The chapter determined that *QTLA* and *QTLB* contribute to the observed variation for DTF by acting on a photoperiod-independent and photoperiod-dependent pathway respectively. Chapter 5 will seek to uncover the molecular basis for *QTLA* and *QTLB*.

Chapter 5 The molecular basis for the control of early flowering in ILL 2601

5.1 Introduction

The genetic basis for the early flowering phenotype of the *pilosae* ILL 2601 was determined in Chapter 4 to be a function of three novel loci. These loci were identified to collectively contribute to the variation in flowering time by regulating three traits, namely *days to flowering* (DTF) and *delay to flower development* (DFD), and *node of floral* initiation (NFI). This chapter explores two of the three identified loci, namely *QTLA* and *QTLB*.

Chapter 4 proposes that *QTLA* regulates the developmental node for NFI and contributes to the observed variation for DTF. *QTLA* is suggested to function independently of prevailing photoperiod, with no difference for NFI observed between inductive and non-inductive photoperiods in ILL 5588. *QTLA* is also described to confer a dominantly inherited early-flowering phenotype. From Chapter 4 is it not certain if a gain-of-function or a loss-of-function is responsible for the early phenotype. This is the first locus in lentil that is reported to control the variation for flowering time independently of photoperiod.

Conversely, *QTLB* was determined to function in the photoperiod-dependent flowering pathway. Chapter 4 proposes that *QTLB* regulates the node delay between NFI and a developed flower, and contributes to the variation for DTF.

This chapter seeks to determine the molecular basis *QTLA* and *QTLB*. The chapter also seeks to determine the prevalence of the candidate for *QTLA* in a collection of lentil accessions.

5.2 Materials and methods

This section details specific materials and methods relevant to this chapter. General materials and methods are described in Chapter 2.

5.2.1 Plant materials and growth conditions

Two F_3 populations derived from F_2 plants 154 and 163 from the ILL 2601 x ILL 5588 F_2 population (Chapter 4) were evaluated under a 12-h short day photoperiod of natural daylength at the University of Tasmania phytotron.

Plant 154 is heterozygous for *QTLA* and homozygous for the ILL 5588 allele at *QTLB* (Figure 5-1). Plant 163 is heterozygous for *QTLB* and homozygous for the ILL 5588 allele at *QTLA* (Figure 5-1).

Poor plant health during early growth and development for both plant 154 and plant 163 F_3 populations had resulted in a high attrition rate, and therefore a small population size (*n* = 20-34).

Forty-seven accessions of cultivated lentil and one *L. culinaris* subsp. *orientalis* accession (ILWL 7) were selected to form a representative collection (Appendix 5) that reflected the diversity of the agro-ecological environment of the regions where lentils are cultivated. The collection was also framed to encapsulate the broad range of flowering times observed for cultivated lentil. Accessions selected were predominantly unimproved landraces (Appendix 5). Accessions from North Africa and the Ethiopian Highlands are underrepresented in the collection. The collection was evaluated under a base photoperiod of 12-h of natural daylength (short day) and supplemented with 4-h of fluorescent lighting (long day) at the University of Tasmania phytotron.

5.2.2 Plant measurements

In this chapter, *days to flowering* (DTF), was measured as first open flower from emergence on either the main stem or the lateral branches. Lateral branches were not excised during DTF evaluation in this chapter.

5.3 Results

5.3.1 Candidate genes for QTLA and QTLB

Genes implicated in flowering time control were annotated on a schematic of lentil linkage group 6 of plant 154 and plant 163. Annotations were made based on their corresponding positions in *M. truncatula* chromosome 7 (Figure 5-1). As described in 5.2.1, F_2 plants 154 and 163 are heterozygous for *QTLA* and *QTLB* respectively. QTL confidence intervals determined in Chapter 4 for *QTLA* and *QTLB* were employed for candidate gene selection. Candidate gene selection will be further elaborated in sections 5.3.2.2 (*QTLA*) and 5.3.3.2 (*QTLB*).

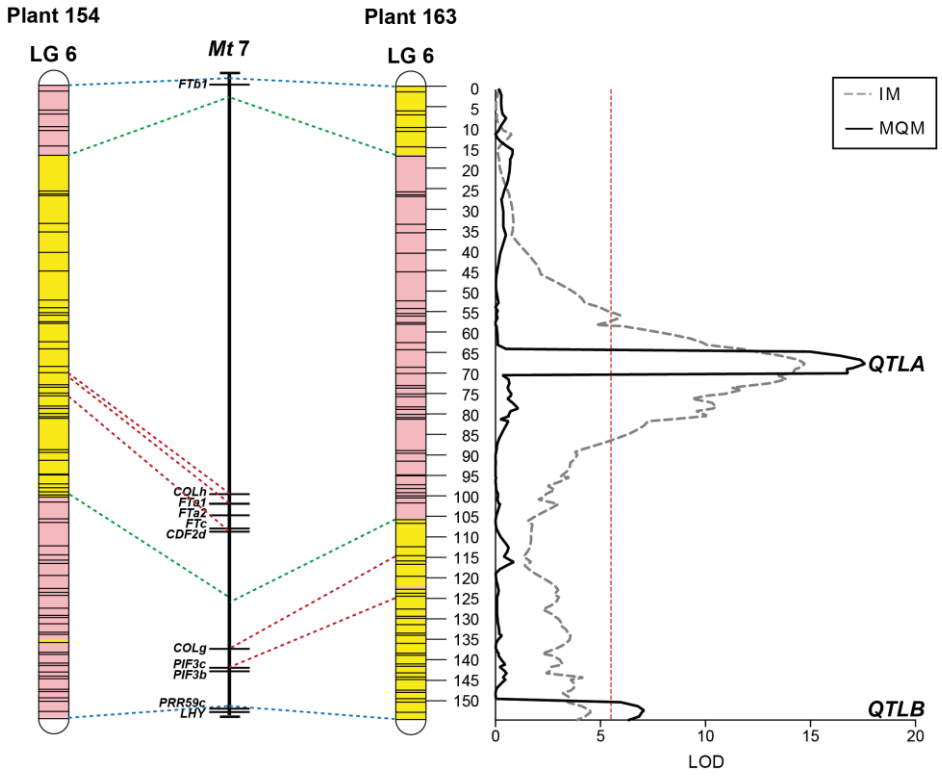


Figure 5-1 Schematic of linkage group 6 for F_2 plants 154 and 163.
Genes associated with flowering time are annotated and regions that are homozygous for a specific parental allele or heterozygous for parental alleles are reflected on a schematic of linkage group 6 for F_2 plants 154 and 163. Regions in yellow represent regions heterozygous for both ILL 5588 and ILL 2601. Regions in pink represent regions homozygous for ILL 5588. In both lines, there were no regions homozygous for ILL 2601. Blue dashed lines represent extent of coverage of *M. truncatula* chromosome 7. Green dashed lines represent boundaries between homozygous and heterozygous regions. Red dashed lines represent approximate position of flowering time genes. QTL graph represent LOD scores for DTF analysed using *Interval Mapping* (IM) and *Multiple-QTL Model mapping* (MQM) (refer to 4.3.6.1 for more details).

5.3.2 Molecular basis for QTLA

In Chapter 4, *QTLA* was determined to account for the largest shift to early flowering in ILL 2601. In order to validate this locus in a genetic context, a F_3 population segregating for *QTLA* was first evaluated under a short day (SD) photoperiod. Candidate genes for this locus were then identified based on their corresponding positions in *M. truncatula*, and their proposed function in other systems. These candidates were subsequently analysed in a time-series expression study under long day (LD) and SD photoperiods. A co-segregation analysis of candidates in a F_3 population was then carried out to test for linkage. To understand the molecular basis for the observed phenotype, the candidate identified was sequenced, and the sequence diversity of the candidate explored across selected accessions. An association of the candidate to both DTF and its prevalence within a collection of lentil accessions was also established.

5.3.2.1 Characterisation of QTLA in F_3 population

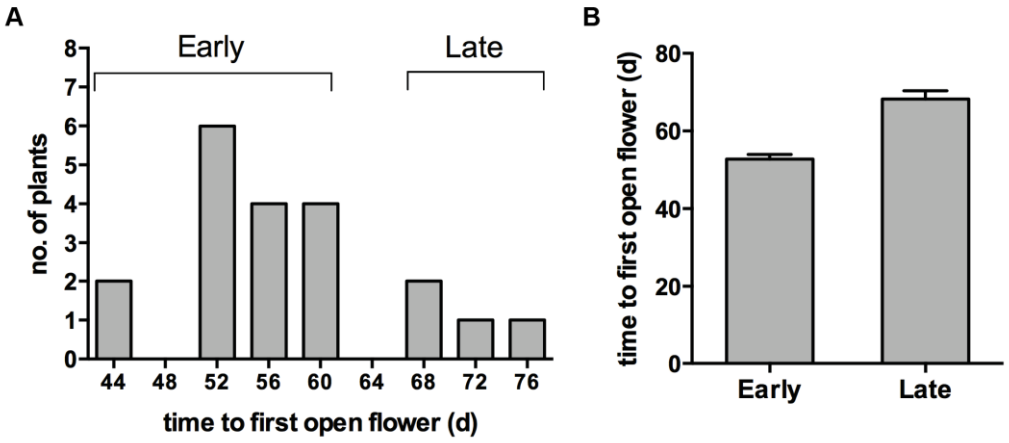


Figure 5-2 Phenotypic characterisation of ILL 2601 x ILL 5588 F_3 population.
(A) F_3 Progeny derived from plant 154 of the ILL 2601 x ILL 5588 F_2 population evaluated under SD conditions for flowering time. Data are for *n*=20. (B) Mean DTF for 'Early' and 'Late' classes. Date are ±SE, n=4-16. Plants received a 12-h photoperiod of natural daylight (SD).

The F_3 progeny (plant 154) were analysed for DTF under SD (Figure 5-2). Bimodality in the segregation was observed with the early class flowering significantly earlier ($p < 0.05$) than the late class. The early class had a mean DTF of 52.8 ± 1.18 days, while the late class had a mean DTF of 68.3 ± 2.14 days.

The 3:1 Mendelian nature (p = 0.606) of the segregation for flowering time points to the dominance of the early flowering phenotype similar to observations in 4.3.6.1.1, and implies that as expected from Figure 5-1, a single locus is segregating for DTF in this population.

5.3.2.2 Candidate gene selection for QTLA

In Chapter 4 it was determined that *QTLA* functions independently of prevailing photoperiod to confer a dominantly inherited early-flowering phenotype. The selection of candidate genes for *QTLA* was therefore restricted to flowering-time genes positioned within the QTL confidence interval that function in or are regulated by pathways independent of photoperiod.

Several flowering-time genes were identified within the confidence interval of *QTLA* (Figure 5-1). Amongst the identified genes, are three *Arabidopsis FLOWERING LOCUS T* (*FT*) orthologues positioned in tandem in *M. truncatula* chromosome 7; namely *FTa1*, *FTa2*, and *FTc*. In *M. truncatula*, *FTa1* and *FTa2* report elevated expression in response to vernalisation (Laurie et al., 2011). *FTa1* is also implicated in the vernalisation-mediated floral induction pathway. Retroelement insertions in the intron or 3' of *FTa1* are described to eliminate the requirement for vernalisation, conferring *M. truncatula* mutants a dominantly inherited early-flowering phenotype (Jaudal et al., 2013; Yeoh et al., 2013). The latter is similar to the early-flowering phenotype conferred by *QTLA*. The photoperiod-independent regulation of *FTa1* and *FTa2* deem them suitable candidates for further analysis. *FTc* is positioned 3' of *FTa2* in *M. truncatula*, and is involved in floral transition (Hecht et al., 2011; Laurie et al., 2011) in *P. sativum* and *M. truncatula*. *FTc* will also be analysed as a candidate.

5.3.2.3 Expression profile of lentil FTa1, FTa2 and FTc

To determine the role of the candidates in flowering time regulation in lentil, a time-series experiment analysing their expression from emergence to one-week post-flowering was carried out (Figure 5-3). In lentil, the expression profiles of *FT* orthologues have not been previously described.

Elevated expression of *FTa1* was observed in ILL 5588 to precede the appearance of flower buds in LD and SD, suggesting that in lentil, as with *P. sativum* (Hecht et al., 2011) and *M. truncatula* (Laurie et al., 2011), *FTa1* functions as a floral promoter (Figure 5-3). *FTa1* expression was also observed to be comparatively higher in the leaf tissue when compared to the apical shoot tissue. In ILL 2601, elevated expression of *FTa1* was observed from emergence in LD and SD. *FTa1* was also expressed at a comparatively higher level during the development of ILL 2601, when compared to ILL 5588. Observations of elevated *FTa1* expression from emergence mirrors *FTa1* upregulation reported in 12-14 day old seedlings of early-flowering *M. truncatula* vernalisation-insensitive mutants (Jaudal et al., 2013; Yeoh et al., 2013).

The expression of *FTa2* was similarly observed to be upregulated from emergence in ILL 2601 (Figure 5-3). However, the relative transcript levels were considerably lower than that observed with *FTa1*. In ILL 5588, *FTa2* expression remained low during the development of the plant in LD and SD, with slight elevated expression observed prior to the appearance of flower buds. Hecht et al. (2011) previously determined that the *P. sativum FTa2* only weakly rescued the late-flowering phenotype of *Arabidopsis ft-1* mutants, suggesting a minor role in floral induction, and likely gene redundancy.

The expression of *FTc* in both ILL 5588 and ILL 2601 was observed to increase with the age (time) (Figure 5-3). However, unlike *FTa1* and *FTa2*, *FTc* was only observed to be upregulated with the appearance of flower buds, suggesting a more downstream role in the flowering pathway. In ILL 2601, elevated expression of *FTc* was not observed, suggesting that *QTLA* is not involved in the direct regulation of the floral promoter. *FTc* is only expressed in the shoot tissue, consistent with findings by Hecht et al (2011) and Laurie et al (2011).

The expression profiles of these candidates during the development of a plant proposes that in ILL 2601, the elevated early expression of both *FTa1* and *FTa2* is associated with early flowering.

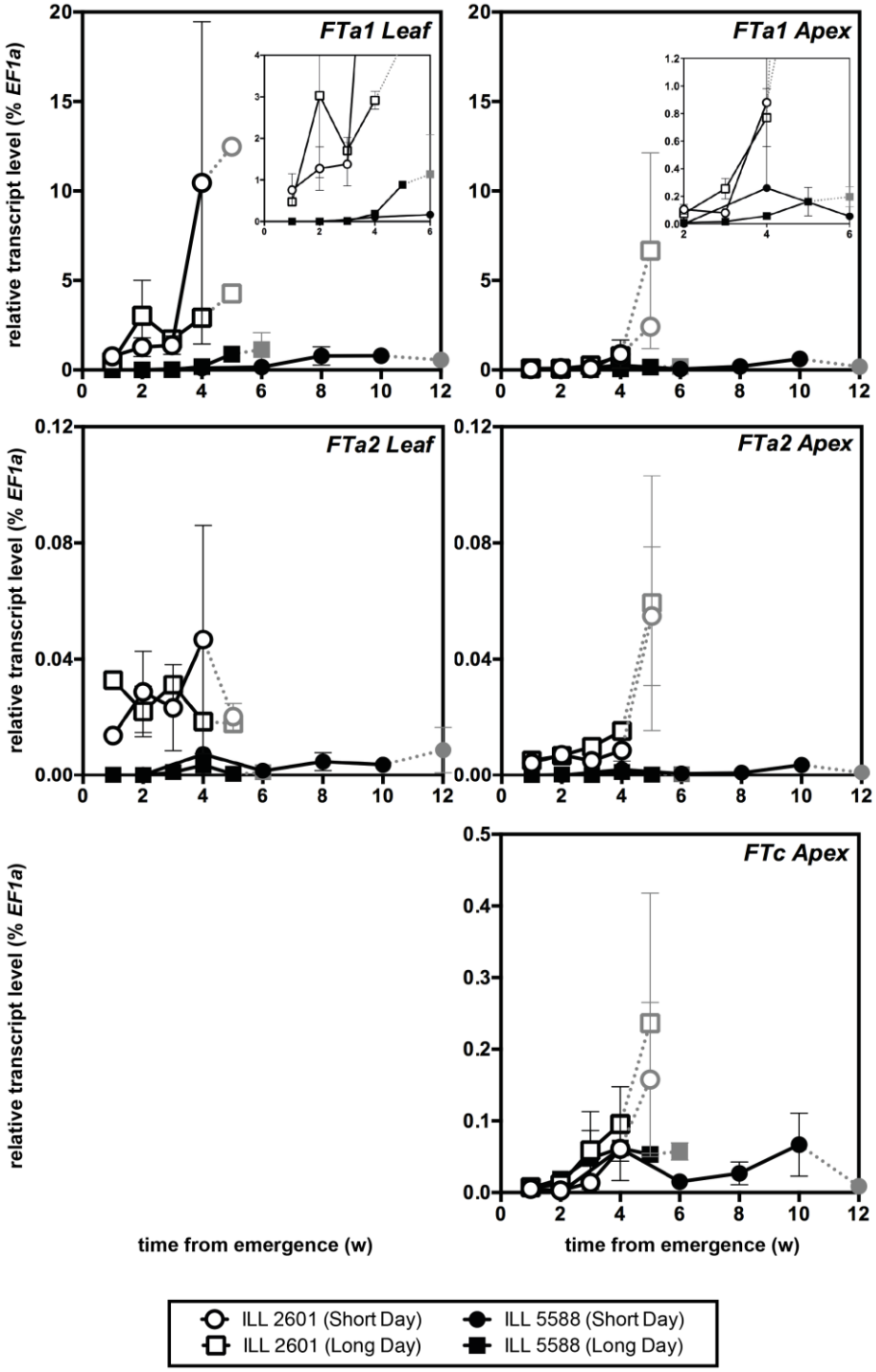


Figure 5-3 Expression of lentil *FT* orthologues under short day and long day photoperiods.
Lentil *FTa1*, *FTa2*, and *FTc* gene expression in photoperiod-sensitive ILL 5588 (closed) and in early-flowering ILL 2601 (open) in dissected shoot apex, and first two fully expanded leaves at one-week intervals from emergence to one-week post flower bud initiation. Plants were exposed to a 12-hour SD photoperiod (circle) and a 16-hour LD photoperiod (square). Grey time-points indicate expression levels post flower bud initiation. Values have been normalised to the transcript level of *ELONGATION FACTOR 1-α* and represent mean ±SE for n=2 biological replicates, each consisting of pooled material from two plants. For *FTa1* initial time-points are magnified inset to show early induction, masked by higher expression during development.

5.3.2.4 Co-segregation analysis for QTLA

To evaluate for linkage between the candidates and the observed variation for flowering time, a co-segregation analysis was carried out with the F_3 progeny from Plant 154 (refer to 5.3.2.1 for flowering time characterisation). The relative distance between *FTa1* and *FTa2*, and the small population size of the F_3 population imply that recombinants between the two *FT* orthologues are unlikely. The population was hence only analysed for *FTa1* co-segregation.

A partial genomic sequence of the lentil *FTa1* orthologue was isolated, and a High-Resolution Melt (HRM) maker designed around a Single Nucleotide Polymorphism (SNP) in the third intron of *FTa1* (refer to Appendix 2 for primer details).

In the co-segregation analysis, it was observed that progeny carrying a single ILL 2601 allele for *FTa1* were early flowering, while progeny homozygous for the ILL 5588 *FTa1* allele were late flowering. These findings point to a tight linkage between *FTa1* and the variation for flowering time, and imply a dominant mode of inheritance for the early flowering phenotype consistent with observations for ILL 2601 alleles at *QTLA* in Chapter 4. The described dominance of the early flowering phenotype is also consistent with the inheritance of vernalisation-insensitivity conferred by retroelement insertions in the intron or 3' of *FTa1* in *M. truncatula* (Jaudal et al., 2013; Yeoh et al., 2013).

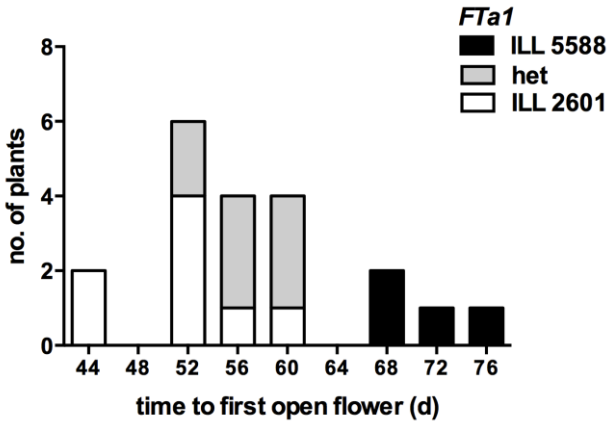


Figure 5-4 Co-segregation of ILL 5588 x ILL 2601 F_3 population for *FTa1* under SD.
F_3 population derived from plant 154 genotyped for *FTa1*. Plants received a 12-h photoperiod of natural daylight (SD). Data are ±SE for *n*=4-8.

However, based on findings by Hecht et al. (2011) and Laurie et al. (2011), it is unlikely that deleterious polymorphisms in the coding region of lentil *FTa1* or *FTa2* that result in a loss of function can lead to the early flowering phenology of ILL 2601. Hecht et al. (2011) determined that *FTa1, FTa2,* and *FTc* act as floral promoters in *P. sativum*, and that deleterious polymorphisms in *FTa1* can result in a late-flowering habit. By inference, it is likely that *FTa1* and *FTa2* are intact and functional in ILL 2601.

5.3.2.5 Isolation and annotation of FTa1-FTa2 cluster

The lentil *FTa1-FTa2* cluster was partially isolated in both ILL 5588 and ILL 2601 using primers designed against sequence information for cv. CDC Redberry provided by Bett, K. (pers. comm. September 2014), and analysed for significant polymorphisms.

A 10335-bp indel 3' of *FTa1* was determined for ILL 2601 in the *FTa1-FTa2* intergenic region (Figure 5-5). In addition, a 2830-bp indel was also reported in ILL 5588 and ILL 2601, positioned 5' of *FTa2*. These observations were made in comparison to sequence information for cv. CDC Redberry (Figure 5-5A).

The *FTa1-FTa2* cluster was subsequently isolated in ILL 4605 (cv. Precoz), ILL 4349 (cv. Laird), cv. Indianhead, and ILWL 7 (*L. culinaris* ssp*. orientalis*) to gain a better appreciation of the diversity for the cluster. ILWL 7, presumably with the wild-type form, was identified to be most similar to ILL 5588 (Figure 5-5A). It was determined that of the accessions studied, the 10335-bp deletion was only present in ILL 2601. ILL 4605 (cv. Precoz), ILL 4349 (cv. Laird), and cv. Indianhead were observed to carry a haplotype that contained the 2830-bp insertion but not the 10335-bp deletion, similar to cv. CDC Redberry (Figure 5-5A).

The absence of the 2830-bp insertion 5' of *FTa2* in ILL 2601 and ILL 5588 rule out the contribution of the insertion to the observed variation for flowering time in this study. BLAST analysis of the sequence for the 2830-bp insertion in LenGen point to the presence of two large transposons in this insertion. The

significance of the 2830-bp insertion 5' of *FTa2* was separately analysed in a F_2 population developed from a cross between cv. Indianhead and ILL 5588. Co-segregation was not observed between the indel and flowering time (LD) (refer to Appendix 6 for flowering time data and analysis). Furthermore, no QTL were determined for this region in Chapter 6 (Table 6-1). Findings collectively rule out any functional significance for this polymorphism in relation to flowering time.

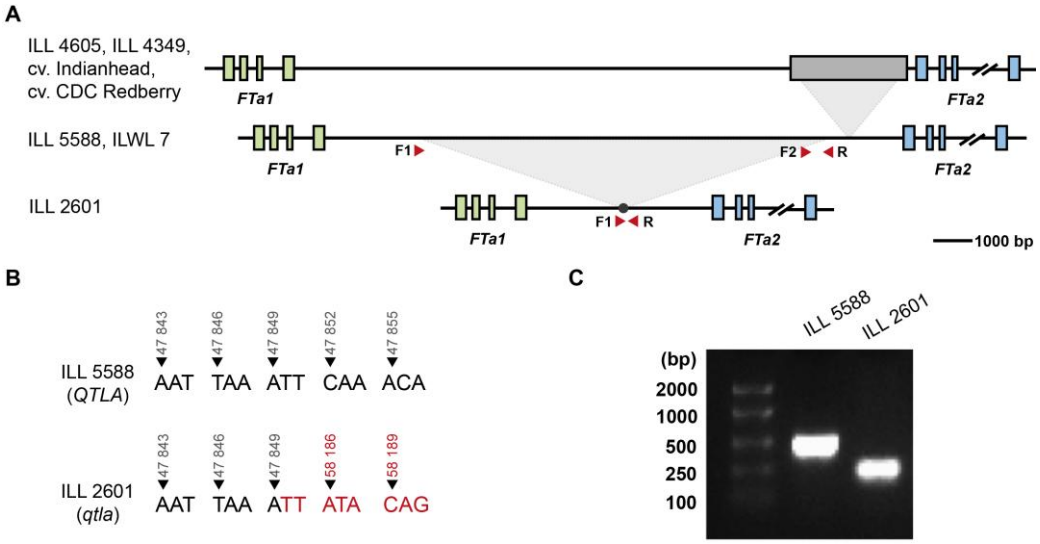


Figure 5-5 Isolation and annotation of the *FTa1-FTa2* cluster.
(A) Schematic diagram of the *FTa1-FTa2* cluster in ILL 5588, ILL 2601, ILL 4605 (cv. Precoz), ILL 4349 (cv. Laird), cv. Indianhead, cv. CDC Redberry, and ILWL 7 (L. culinaris ssp. orientalis). The green boxes represent the exons of the lentil FTa1 and the blue boxes represent the exons of lentil FTa2. The grey box represents a 2830-kb indel in the 5' of FTa2. (B) The sequence position of the 10335-bp deletion in ILL 2601. (C) PCR of the intergenic region with a 45s extension time (approximately 1-Kb) in ILL 5588 and ILL 2601. Primers positions are annotated in (A). Refer to Appendix 3 for primer details for sequence isolation and Appendix 2 for FTa1-FTa2 allele-specific PCR marker details.

5.3.2.6 Transcript profile of FTa1-FTa2 cluster

The presence of retroelements 3' of *FTa1* in *M. truncatula* has been proposed to reduced the requirement for vernalisation, and confer an early-flowering phenology (Jaudal et al., 2013). Jaudal et al. (2013) also suggested that retroelements in the *M. truncatula* mutants may interfere with the regulatory elements involved in vernalisation. In ILL 2601, the deletion of 10335-bp of non-coding sequence is likely to invoke a similar disruption to the regulation of the vernalisation response. However, the non-coding *FTa1-FTa2* intergenic region has not been described in lentil, or in members of the *Fabeae* tribe.

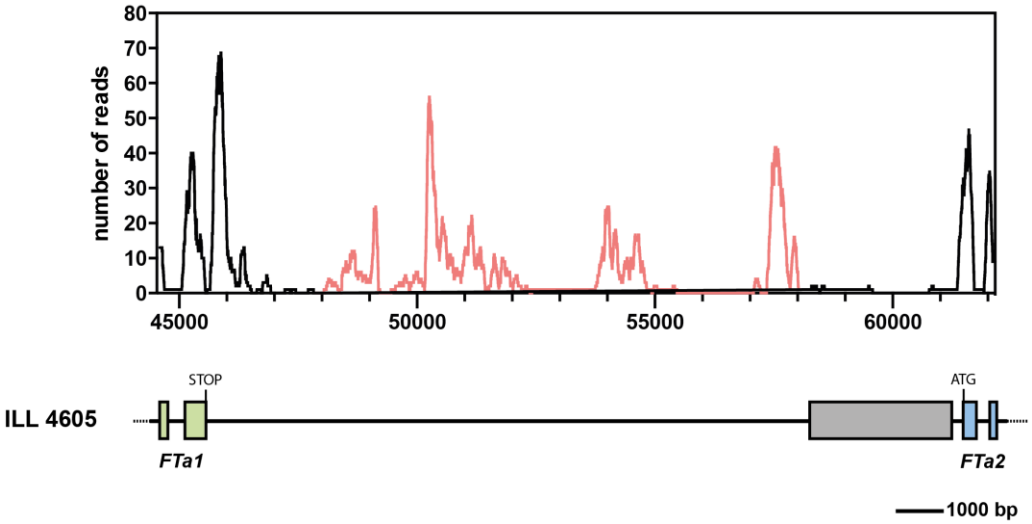


Figure 5-6 Transcript profile of *FTa1-FTa2* intergenic region in lentil accession ILL 4605 (cv. Precoz).
The red graph plot represents the 10335-bp deleted interval in ILL 2601. The green boxes represent exons 3 and 4 of the lentil *FTa1* and the blue boxes represent exons 1 and 2 of lentil *FTa2*. The grey box represents a 2830-kb indel in the 5' of *FTa2* consisting of two transposons, observed in several lines. The 2830-kb indel was excluded from the analysis.

The transcript profile in ILL 4605 for the deleted intergenic region between *FTa1-FTa2* was analysed using data from the Sequence Read Experiment SRX31720 of the Sequence Read Project SRP026548. The publicly available sequence reads from SRP026548 were originally intended for an experiment aimed at SNP discovery for the construction of a genetic map for a cv. Precoz x WA8649041 RIL population segregating for the lentil *Sn* (Kahriman et al., 2014). The transcript data for ILL 4605 is hosted on http://www.ncbi.nlm.nih.gov/sra and is the only publicly available transcriptome resource for a University of Tasmania in-house lentil accession.

A relatively high level of reads was observed in ILL 4605 for the *FTa1-FTa2* intergenic region, within the 10335-bp deletion reported in ILL 2601 (Figure 5-6). This suggests that the non-coding sequence in this region is expressed in lentil, alluding to the potential presence of non-coding RNA (ncRNA) in the *FTa1-FTa2* intergenic region. The 2830-kb insertion 3' of *FTa2* was excluded from the analysis.

5.3.2.7 Effect of FTa1-FTa2 deletion on flowering time

A co-segregation analysis was carried out to determine the association of the *FTa1-FTa2* 10335-bp deletion to the observed early-flowering phenotype. As expected from section 5.3.2.4, it was observed that progeny carrying the ILL 2601 haplotype for the *FTa1-FTa2* intergenic region were entirely early flowering, with progeny homozygous for the ILL 5588 haplotype reportedly later flowering (Figure 5-7A). F_3 progeny homozygous for the ILL 2601 haplotype were reported a mean DTF of 50.5 ± 1.88 days, and was not observed to be significantly earlier (p = 0.0547) than the heterozygous class (mean DTF = 55.0 ± 1.04 days), suggesting a dominant mode of inheritance (Figure 5-7B). This is consistent with the characterisation of *QTLA* in Chapter 4, and observations by Jaudal et al. (2013) and Yeoh et al. (2013) in *M. truncatula*. F_3 progeny homozygous for the ILL 5588 haplotype flowered significantly later than both classes, with a reported mean DTF of 68.3 ± 2.14 days (Figure 5-7B).

An association analysis was carried out to understand the contribution of the 10335-bp deletion reported in the *FTa1-FTa2* intergenic region to the observed variation for DTF in a representative collection of forty-eight lentil accessions selected for their diverse geographic origins evaluated under LD and SD. The lentil accessions demonstrated a wide continuous variation for flowering time (Figure 5-7C). Refer to 5.2.1 for more information pertaining to the composition of the collection.

It was observed that within the collection, all thirteen accessions sourced from countries in the Indian Subcontinent were the earliest to flower in SD, with the exception of accessions homozygous recessive for the lentil *Sn* (*elf3-1*) (Figure 5-7C). In addition, twelve of the thirteen accessions from the Indian Subcontinent evaluated, reported the *FTa1-FTa2* 10335-bp deletion (Figure 5-7C). The only exception was accession PI 426797, which is of Pakistani origin.

However, the prevalence of the *FTa1-FTa2* 10335-bp deletion was not exclusive to accessions from the Indian Subcontinent (Figure 5-7D). Instead, six other

accessions, one each from Lebanon (ILL 2276), Afghanistan (ILL 1823), Turkey (PI 339293) and Tajikistan (PI 606610), and two from Iraq (ILL 2153, ILL 4370), also reported the *FTa1-FTa2* 10335-bp deletion (Figure 5-7C and Figure 5-7D). These accessions were later flowering in SD, when compared to the accessions from the Indian Subcontinent (Figure 5-7C). Furthermore, the accession from Tajikistan (PI 606610) was reported to not flower within 140 days, under SD conditions (Figure 5-7C).

The incidence of the early-flowering ILL 2601 haplotype in the Afghan and Tajik accessions present an interesting perspective for lentil adaptation to the higher latitudes.

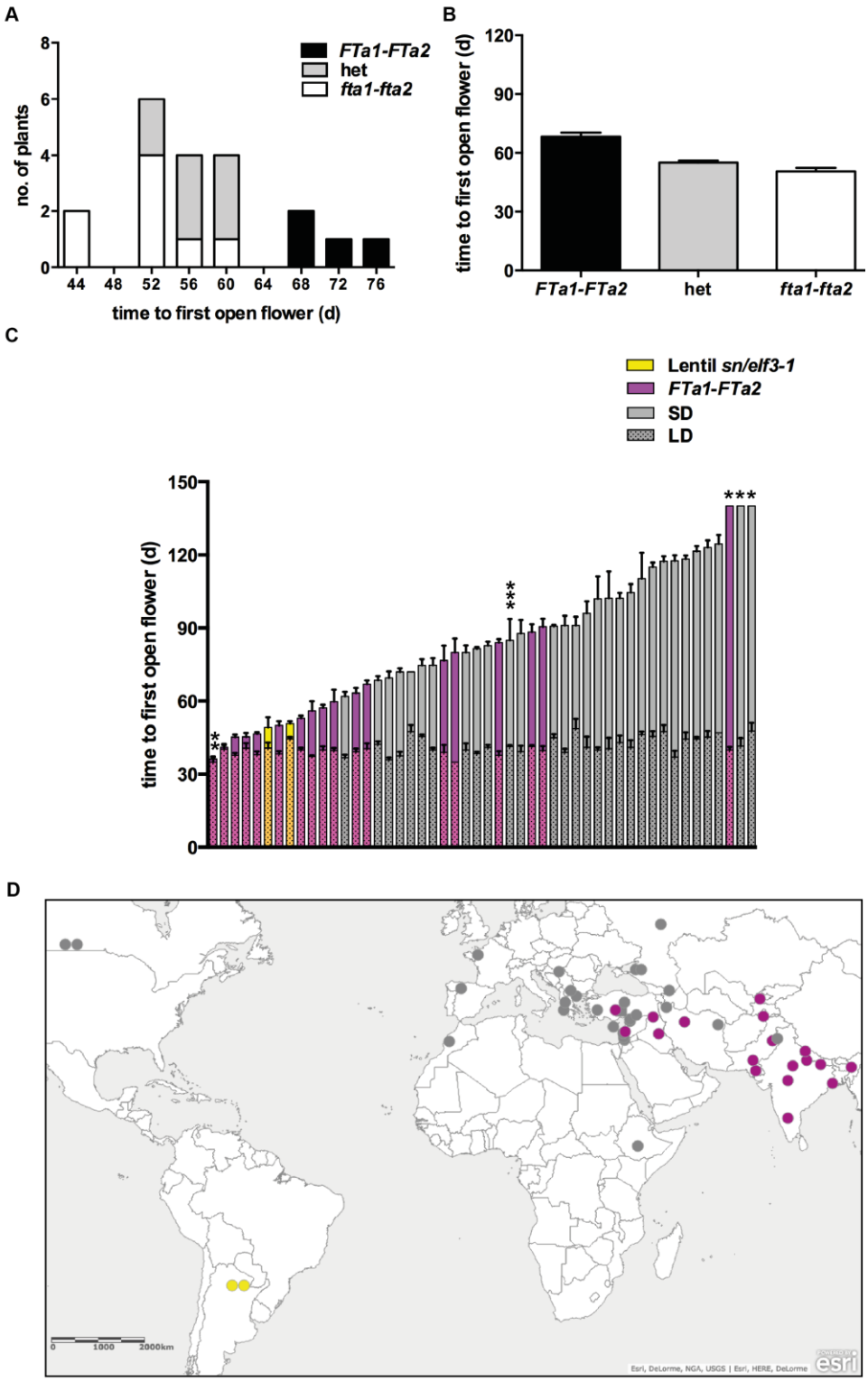


Figure 5-7 Association analysis of *FTa1-FTa2* deletion and flowering time.
(A) F_3 population derived from plant 154 evaluated under SD condition for flowering time and genotyped for *FTa1-FTa2* deletion. Data are for *n*=4-8. (B) Mean DTF for each haplotype in the ILL 2601 x ILL 5588 F_3 population under SD conditions. Data are ±SE for *n*=4-8. (C) DTF of University of Tasmania in-house lentil accessions under LD and SD. (*) denotes accessions that did not flower after 140 days in SD. (**) denotes ILL 2601. (***) denotes ILL 5588. (D) Prevalence of *FTa1-FTa2* haplotypes. Grey circles indicate wild-type (ILWL 7/ILL 5588) haplotype, Purple circles indicate ILL 2601 haplotype (deletion), and yellow circles indicate *elf3-1*.

5.3.3 Molecular basis for QTLB

In Chapter 4, *QTLB* is proposed to regulate DFD and contribute to the variation for DTF. DFD is photoperiod responsive, and it was hypothesised that *QTLB* functioned in the photoperiodic pathway.

In this section, the molecular basis for *QTLB* is investigated using candidate-gene mapping, co-segregation analysis of candidate for DTF in a F_3 population derived from ILL 2601 x ILL 5588 F_2 progeny heterozygous for *QTLB*, and validated through sequence analysis.

5.3.3.1 Characterisation of QTLB in F_3 population

Progeny from plant 163 of the ILL 2601 x ILL 5588 F_2 population were analysed for DTF under SD. F_2 plant 163 is heterozygous for *QTLB* and homozygous for the ILL 5588 allele at *QTLA*. While a large variation in DTF was observed in the F_3 population, distinct classes of late and early segregants were not reported.

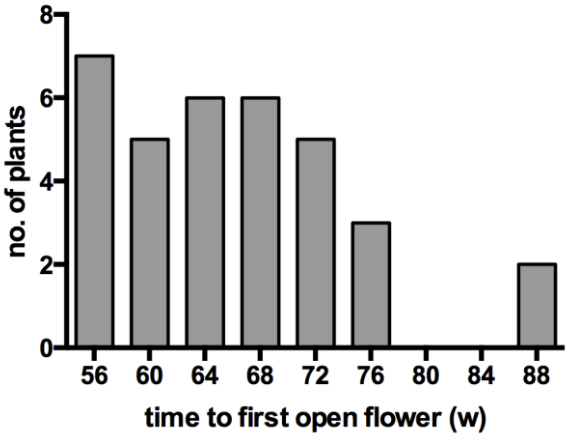


Figure 5-8 Phenotypic characterisation of ILL 2601 x ILL 5588 F_3 population.
F_3 progeny derived from plant 163 of the ILL 2601 x ILL 5588 F_2 population evaluated under SD conditions for flowering time. Data are for *n*=34. Plants received a 12-h photoperiod of natural daylight.

5.3.3.2 Co-segregation and mapping of candidate genes

A broad QTL peak was determined in Chapter 4 for *QTLB* during interval mapping for DTF in the ILL 2601 x ILL 5588 F_2 population. This peak was further defined through MQM mapping. However, the QTL peak was observed to occur at the end of the defined linkage group 6, and significant LOD scores for all

markers at the end of were reported (Figure 4-6). Furthermore, it was also observed that the lentil linkage group 6 did not afford full coverage of the corresponding region in *M*. *truncatula* chromosome 7 for *QTLB* (Figure 5-1 and Figure 4-5). This suggests that it is likely that the *QTLB* interval for DTF and DFD extends beyond the defined lentil linkage group 6. It is also plausible that the candidate of interest is not positioned within the defined linkage group. It was therefore necessary to further define the end of the lentil linkage group 6 by mapping.

Arabidopsis homologues of four genes with predicted roles in the photoperiodic pathway were identified in *M*. *truncatula* chromosome 7 (Table 5-1). These genes are positioned at the end of *M*. *truncatula* chromosome 7, which corresponds to lentil linkage group 6 with elevated LOD scores for DTF and DFD. Partial genomic sequences of the lentil orthologues for these genes were isolated (refer to Appendix 3 for primer information) in ILL 5588 and ILL 2601, and makers were designed around polymorphisms (refer to Appendix 2 for marker information) in the parental alleles of these genes.

Gene symbol	Medicago locus (Mt4.0)	Gene description	Additional information
COLg	Medtr7g108150	zinc finger *CONSTANS*-like protein	weak diurnal expression rhythm in LD and SD, unlikely to function as central integrator of photoperiod responsive flowering in M. truncatula (Wong et al., 2014).
PIF3c	Medtr7g110810	helix loop helix DNA-binding domain protein	*Arabidopsis PIF3* orthologue is diurnally regulated, and mediates light dependent growth (Soy et al., 2014).
PRR59c	Medtr7g118260	Pseudo-Response Regulator	*P. sativum PRR59a* paralogue demonstrates strong diurnal expression rhythm in LD and SD (Liew et al., 2009a). *Arabidopsis PRR* mutants demonstrate defects in photoperiod control of flowering time (Nakamichi et al., 2007).
MYB1/ LHY	Medtr7g118330	late elongated hypocotyl-like protein	*P. sativum MYB1* demonstrates strong diurnal expression rhythm in LD and SD. *P. sativum MYB1* is a *Arabidopsis CCA1/LHY* orthologue (Hecht et al., 2007; Liew et al., 2009a).

Table 5-1 Genes with predicted roles in the photoperiodic pathway.
Partial genomic sequence of the lentil orthologues for *COLg*, *PIF3c*, *PRR59c*, and *MYB1* were isolated in ILL 5588 and ILL 2601, and makers were designed around polymorphisms in the parental alleles of these genes.

The lentil orthologues were mapped in the ILL 2601 x ILL 5588 F_3 population derived from F_2 plant 163. While the orthologues were observed to map approximately to the expected corresponding *M. truncatula* positions, a discrepancy in the order of *PRR59c* and *MYB1* was reported (Figure 5-9).

QTL mapping was carried out to determine the contribution of these orthologues to the reported DTF in the F_3 population. While *LcPRR59c*, *LcMYB1*, and *LcPIF3c* reported elevated LOD scores when analysed using interval mapping, it was determined through MQM mapping that *LcPRR59c* is the most likely candidate. *LcPRR59c* reported a LOD score of 3.22, and is estimated to account for 35.3% of the observed variation in DTF in the F_3 population. The reported LOD score was however below the LOD significance threshold of 4.1.

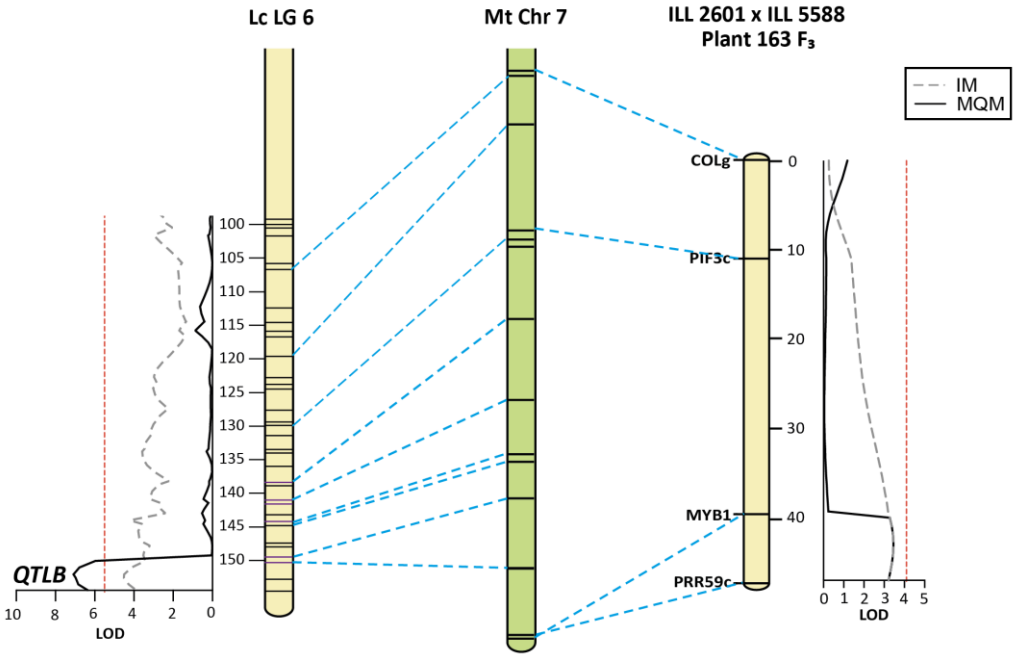


Figure 5-9 Relationship between lentil linkage group 6, *M. truncatula* chromosome 7, and mapped lentil orthologues in ILL 2601 x ILL 5588 F_3 population.
Intervals in lentil linkage maps indicate cM distances between loci, estimated from segregation data using JoinMap 4.0 (Van Ooijen, 2006). The physical map for M. truncatula chromosome 7 is based on Mt4.0 Medicago reference genome (Tang et al., 2014). Where appropriate, orthologues that correspond between maps are linked with dotted blue lines. The LOD scores indicate contribution of mapped loci to observed variation in DTF. LOD threshold (dotted red) is 5.5 for Lc LG 6 in the ILL 2601 x ILL 5588 F_2 population and 4.1 in the ILL 2601 x ILL 5588 F_3 population.

5.3.3.3 Candidate gene identification and association analysis with QTLB

In the identification of candidate genes with a predicted role in the photoperiodic pathway for *QTLB*, positioned within the corresponding region in *M. truncatula*, it was determined that only the *PRR59c* and *MYB1* orthologues had the potential to explain the observed variation in DTF. *MYB1* was excluded as a candidate through QTL mapping, where it was determined that the orthologue only explained an estimated 2% of phenotypic variation.

A co-segregation analysis was undertaken to determine the contribution of *LcPRR95c* to the observed variation in DTF. The alleles of either parent were not observed to complete co-segregate for DTF (Figure 5-10). However, progeny homozygous for the ILL 2601 allele were reported to flower significantly earlier ($p < 0.05$) than progeny carrying a single allele from the ILL 5588 parent (Figure 5-10). Heterozygote progeny were observed to be an intermediate between both homozygous classes (Figure 5-10). A co-segregation analysis was also undertaken to determine the contribution of *LcMYB1* to the observed variation in DTF (refer to Appendix 7).

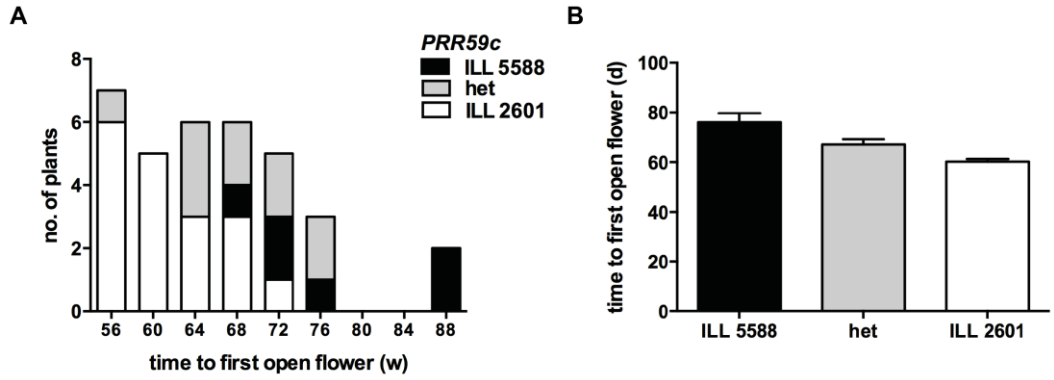


Figure 5-10 Co-segregation of ILL 5588 x ILL 2601 F_3 population with lentil *PRR59c* under SD. (A) Progeny from Plant 163 of the ILL 2601 x ILL 5588 F_2 population evaluated under SD conditions for flowering time. (B) Mean DTF for progeny for each allele in the ILL 2601 x ILL 5588 Plant 163 F_3 population under SD conditions. Data are ±SE for *n*=6-18.

5.3.3.4 Annotation of lentil PRR95c

The *PSEUDO-RESPONSE REGULATOR* (*PRR*) gene family has been extensively characterised in *Arabidopsis*. The current literature proposes that genes within the *PRR* family function to promote flowering, and that recessive mutations

confer a late-flowering phenology (Matsushika et al., 2002; Nakamichi et al., 2012; Nakamichi et al., 2005; Sato et al., 2002).

The full-length genomic and coding sequences of the lentil *PRR59c* were isolated, and sequenced in both ILL 5588 and ILL 2601. In the analysis for polymorphisms, an insertion-deletion (indel) in exon 8 was identified (Figure 5-11). The presence of a single guanine (G) residue in the ILL 2601 allele of the lentil *PRR59c* results in a frame-shift, and a predicted premature stop codon in the early-flowering parent. The frame-shift incidentally occurs within the conserved CO, CO-like, and TOC1 (CCT) domain.

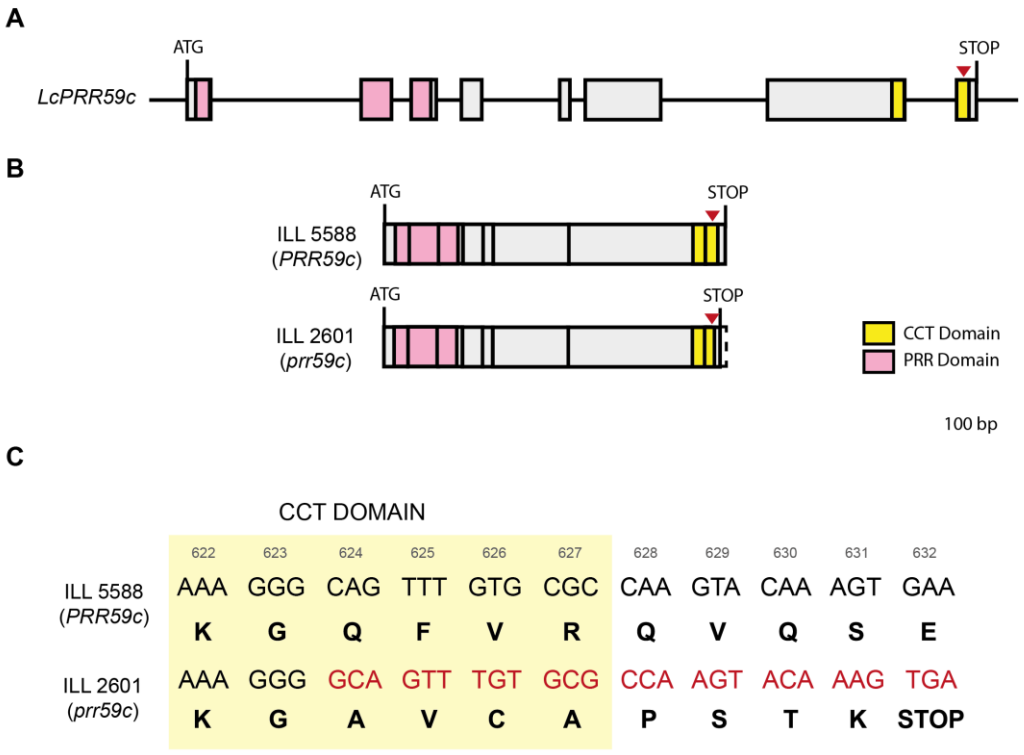


Figure 5-11 Lentil *PRR59c* and nature of polymorphism in early-flowering ILL 2601.
(A) Schematic of lentil *PRR59c* PRR and CCT domains, and location of polymorphism (red arrow) in ILL 2601. (B) Schematic of predicted frame-shift and premature stop in the ILL 2601 *prr59c* mRNA. (C) Details of the predicted truncation of ILL 2601 PRR59c protein.

To further probe the significance of the frame-shift mutation, the CCT domain across selected legumes, *Oryza sativa, Brachypodium distachyon, Arabidopsis thaliana,* and *Populus trichocarpa* was analysed (refer to Appendix 4 for

sequence information). The sequence analysis determined that the ILL 5588 allele is the functional allele for the lentil *PRR59c* (Figure 5-12). The lentil *PRR59c* has not been previously reported in the literature. In legumes, only one orthologue from the PRR5/9 clade has been previously reported (Liew et al., 2009b). Described in *P. sativum* and designated *PsPRR59*, the legume PRR gene is determined to be orthologous to *Arabidopsis PRR5* (Liew et al., 2009b), and is described to be diurnally regulated with expression peaking during the night (Liew et al., 2009a). In this thesis, the legume *PRR59* is provisionally reassigned *PRR59a*.

To establish the relationship of the *PRR* orthologue to members of the PRR5/9 clade, a phylogenetic analysis that included selected legumes*, O. sativa, B. distachyon, A. thaliana,* and *P. trichocarpa* was carried out (Figure 5-13 and Appendix 8).

It was determined that two major groups exist within the PRR5/9 clade of dicots; each demonstrating sequence homology with Arabidopsis *PRR5* and *PRR9* respectively (Figure 5-13 and Appendix 8). In monocots, the duplication of the PRR5/9 ancestor occurred after the speciation of monocots and dicots (Takata et al., 2010). As illustrated in Figure 5-13, it was determined that *PRR59c* is likely an *Arabidopsis PRR9* orthologue. Furthermore, a legume-specific duplication of the *Arabidopsis PRR9* orthologue is also proposed (Figure 5-13). The paralogues in legumes are provisionally assigned *PRR59b* and *PRR59c* respectively (Figure 5-13). The role of the legume-specific *PRR59* gene family in floral induction is not defined in the current literature.

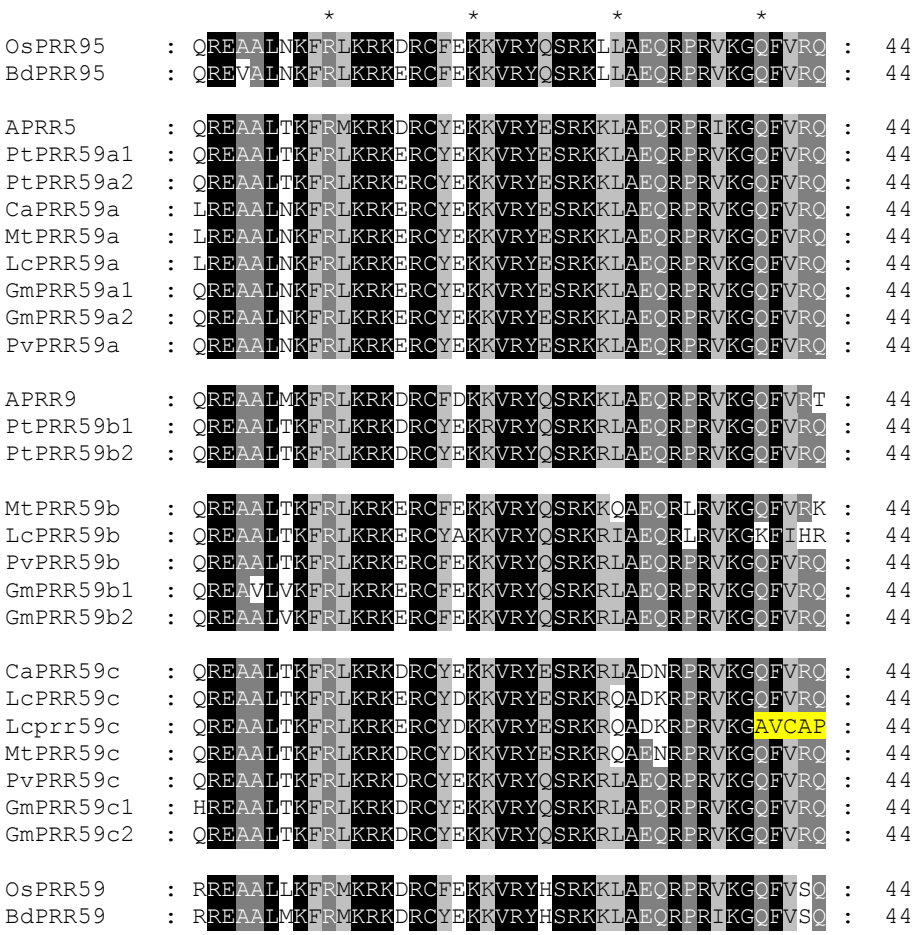


Figure 5-12 Conservation of CCT domain across PRR homologues.
The alignment was created with predicted protein sequences for the CCT domain of selected legumes*; O. sativa* (Os)*, B. distachyon* (Bd)*, A. thaliana* (At)*,* and *P. trichocarpa* (Pt) aligned with ClustalX and manually adjusted and annotated using GeneDoc and Adobe Illustrator. Shading indicates degrees of conservation; black=100%, dark grey=80%, light grey=60%, yellow=frame-shift. Species abbreviations are as follows: *Lens culinaris* (Lc), *Pisum sativum* (Ps), *Medicago truncatula* (Mt), *Cicer arietinum* (Ca), *Glycine max* (Gm), *Phaseolus vulgaris* (Pv). Lcprr59c is the LcPRR59c predicted protein from ILL 2601. Refer to Appendix 4 for sequence information.

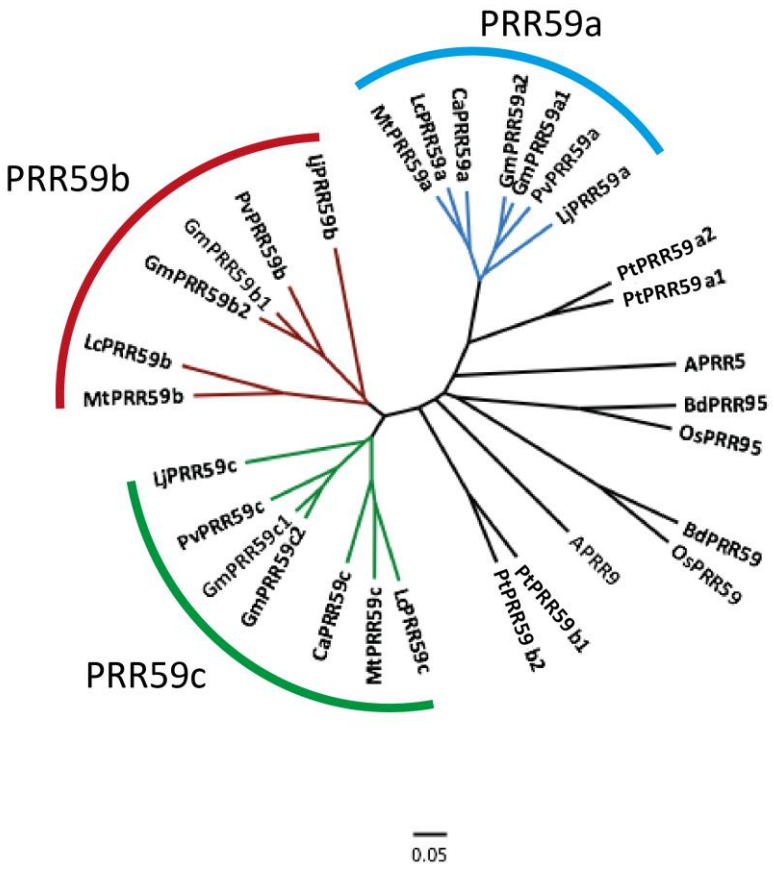


Figure 5-13 Phylogenetic relationship of PRR5/9 Clade.
Phylogram was created using full-length protein sequences of selected legumes, *O. sativa, B. distachyon, A. thaliana,* and *P. trichocarpa* aligned with ClustalX and visualised using FigTree v1.4.2. Duplicated legume PRR59 paralogues are assigned PRR59b and PRR59c respectively. Species abbreviations are as follows: *L. culinaris* (Lc), *Lotus japonicas* (Lj) *M. truncatula* (Mt), *C. arietinum* (Ca), *G. max* (Gm), *P. vulgaris* (Pv), *A. thaliana* (At). *O. sativa* (Os), *B. distachyon* (Bd), and *P.* trichocarpa (Pt). Refer to Appendix 4 for sequence information and Appendix 8 for full alignment.

5.4 Discussion

The current understanding of the molecular basis for the control of early flowering in lentil is largely limited to the study of the lentil *Sn*. In Chapter 4, evidence is provided that three novel loci are also involved in the control of flowering in lentil. These loci are determined to collectively contribute to the early flowering phenotype of ILL 2601. This chapter builds on these findings by characterising the molecular basis for the flowering time loci *QTLA* and *QTLB*.

5.4.1 Molecular basis for QTLA

The flowering locus *QTLA* is established in Chapter 4 to be involved in the regulation of DTF and the developmental node for NFI. It is proposed that the locus is dominantly inherited and regulated independently of photoperiod. *QTLA* occurs in the corresponding *M. truncatula* region that encompasses the *Arabidopsis FT* homologues, *FTa1, FTa2, and FTc*.

The *Arabidopsis FT* is a member of the phosphatidylethanolamine binding domain protein (PEBP) family that promotes flowering through its interaction with CONSTANS (CO) (Kardailsky et al., 1999; Kobayashi et al., 1999). In legumes, six homologues of the *Arabidopsis FT* have been previously reported (Hecht et al., 2011; Weller and Ortega Martinez, 2015), three of which are located within the QTL confidence interval for *QTLA.*

In lentil, the roles of the *Arabidopsis FT* orthologues in the control of flowering time have not been previously reported. Through the study of the gene expression profiles of *FTa1*, *FTa2*, and *FTc* from emergence to flowering in this chapter, it can be suggested that in lentil the upregulation of all three *FT* homologues are associated with the appearance of flower buds, alluding to their potential role in the lentil flowering pathway. Additionally, in early-flowering ILL 2601, it was observed that both *FTa1* and *FTa2* are significantly upregulated one-week from emergence, albeit at comparatively lower levels with the latter. This is in contrast to the expression profiles for these genes

observed in ILL 5588, where *FTa1* and *FTa2* are only upregulated prior to the appearance of flower buds, later in the development of the plant.

In the phylogenetically related *P. sativum*, all three *Arabidopsis FT* orthologues have been proposed to function as floral promoters, with functional copies of the *P. sativum* homologues established in a complementation study by Hecht et al. (2011) to rescue the late-flowering phenology of the *Arabidopsis ft-1* mutant. Hecht et al. (2011) and Laurie et al. (2011) additionally determined that in *P. sativum* and in *M. truncatula*, deleterious polymorphisms within the coding region of *FTa1* results in a late-flowering phenology, suggesting that for early flowering its is likely the *FT* orthologues in lentil are intact.

Observations in *P. sativum*, coupled with the reported expression profiles of all three *Arabidopsis FT* homologues in lentil suggests that it is plausible that in lentil, *FTa1*, *FTa2*, and *FTc* function to promote flowering. Additionally, the increased expression of both *FTa1*, *FTa2* in ILL 2601 one-week from emergence suggests a specific role for these genes in the promotion of early flowering in ILL 2601.

Based on the proposed role for the lentil *FT* orthologues, it is not conceivable that a deleterious polymorphism in the coding sequence of the either lentil *FTa1* or *FTa2* can result in a gain-of-function phenology as observed in the early-flowering landrace ILL 2601. Instead, it can be hypothesised the lentil *QTLA* functions to promote flowering through the upregulation of both *FTa1* and *FTa2*.

Through partial isolation and sequencing of the lentil *FTa1* and *FTa2* cluster in this chapter, a 10335-bp deletion in the *FTa1-FTa2* intergenic region, 3’ of the lentil *FTa1* orthologue was revealed in the early-flowering ILL 2601. This deletion is also surveyed to be prevalent across most South Asian accessions, and to occur in the Afghan and Tajik accessions surveyed. Additionally, an analysis of the transcript profile of ILL 4605 for this region suggests that in lentil

the non-coding intergenic region is expressed, and that a single or potentially multiple ncRNA maybe present. Neither sequence deletions nor the role of ncRNA within the *FTa1-FTa2* intergenic region have not been previously implicated in the control of flowering time.

Incidentally in *M. truncatula*, retroelement insertions in or 3' of *FTa1* have been described to promote flowering, conferring mutants an early-flowering phenology (Jaudal et al., 2013; Laurie et al., 2011). Furthermore, it is proposed that the region 3' of *FTa1* is involved in the control and regulation of the vernalisation response, where the retroelement insertions eliminate the requirement for vernalisation and results in the upregulation of *FTa1* from germination (Jaudal et al., 2013). Jaudal et al. (2013) further proposes that the early-flowering habit conferred by these retroelement insertions is dominantly inherited.

In lentil, the early-flowering habit conferred by *QTLA* is dominantly inherited and regulated independently of photoperiod, consistent with observations in *M. truncatula*. Additionally, in the early-flowering ILL 2601, it was observed that both *FTa1* and *FTa2* are significantly upregulated one-week from emergence, suggesting a possible association of the 10335-bp deletion with the upregulation of both floral promoters, again consistent with observations of *FTa1* upregulation in early-flowering *M. truncatula* vernalisation mutants (Jaudal et al., 2013). Furthermore, it a population segregating for *QTLA* the early-flowering habit co-segregated completely with the 10335-bp deletion, suggesting that this mutation may be causal.

ncRNA have been previously implicated in various regulatory processes in both plant and animal systems. In plants ncRNA have been implicated in numerous processes including the regulation of developmental process such as the transition to flowering (Liu et al., 2015). In *Arabidopsis*, two long non-coding RNA (lncRNA) have been implicated in the regulation of the vernalisation response. It is suggested that the long intronic non-coding RNA, designated

COLDAIR (*COLD-ASSISTED INTRONIC NON-CODING RNA*) (Heo and Sung, 2011), and a long antisense RNA, designated *COOLAIR* (*COLD INDUCED LONG ANTISENSE INTRAGENIC RNA*) (Swiezewski et al., 2009) are necessary for the vernalisation-mediated repression of the floral repressor *FLOWERING LOCUS C* (*FLC*) MADS-box gene. In legumes however, it is suggested that orthologues of *FLC* do not exist (Hecht et al., 2005). Moreover, in *Arabidopsis* the lncRNA functions to promote flowering by repressing a floral inhibitor, distinct from the upregulation of floral promoters *FTa1* and *FTa2* in lentil.

While the *Arabidopsis* vernalisation model is distinct from the anticipated role of *QTLA* in lentil, alternative mechanisms which detail the regulation of promoter genes by ncRNA have been described in other systems. One possibility is that the *FTa1-FTa2* non-coding intergenic region lentil functions to promote and inhibit flowering by alternating between its cis-trans isoforms. A similar mechanism has been described in mice for the regulation of both *Dlx5* and *Dlx6*, involved in appendicular skeletal development (Berghoff et al., 2013; Bond et al., 2009; Feng et al., 2006). These genes are regulated by a lncRNA designated *Evf2* positioned in the intergenic region between the two genes. Upon knockout of the *Evf2* transcript, an upregulation of both *Dlx5* and *Dlx6* is described (Bond et al., 2009). *Evf2* is the first ncRNA reported to function both as an enhancer and repressor in any system (Berghoff et al., 2013; Feng et al., 2006).

More work is required to further probe the role of the non-coding intergenic region between *FTa1* and *FTa2* in conferring ILL 2601 an early flowering phenology, and the role of transcribed ncRNA from this region in the photoperiod-independent flowering induction pathway in lentil.

Collectively, work from this chapter point to a single or potentially multiple ncRNAs, present in the non-coding *FTa1-FTa2* intergenic region, of having a function in regulating the promotion of *FTa1* and *FTa2* in lentil. It is plausible that the retroelement insertions reported in the *M. truncatula* vernalisation mutants (Jaudal et al., 2013) function to disrupt the ncRNA, hence impairing its

function as an inhibitor of *FTa1.* Additionally, it can be proposed that the regulation, or the potential absence of regulation of *FTa1* and *FTa2* in ILL 2601 results in the high levels of expression observed with the floral promoters resulting in early flowering, similar to observations of *Dlx5* and *Dlx6* in mice. Furthermore, based on the relative transcript levels of both *FTa1* and *FTa2* expression, it is likely that in lentil both *FTa1* and *FTa2* demonstrate redundancy, with *FTa1* having a greater effect on floral promotion. The minor role of the lentil *FTa2* suggested is supported by previous work in *P. sativum* where it was established that a functional *FTa2* only weakly rescues the *Arabidopsis ft-1* mutant phenotype (Hecht et al., 2011).

This chapter does not explore vernalisation, and cannot directly attribute the early-flowering phenology to vernalisation-insensitivity. Additionally, how the ncRNA responds to vernalisation is not known. These need to be further explored. The molecular basis for the vernalisation response is not known in lentil or in other legumes.

5.4.2 Contribution of lentil FTa1-FTa2 intergenic region to adaptation and spread

The contribution of the 10335-bp deletion in the non-coding *FTa1-FTa2* intergenic region to the adaption of lentils, and its cultivation in the Indo-Gangetic plain is both interesting from a crop adaptation, and a plant breeding perspective.

Lentils in Afghanistan are spring-sown. Incidentally, accessions from Afghanistan and neighbouring Tajikistan, while amongst the latest to flower in short day conditions, carry the *FTa1-FTa2* intergenic deletion. Erskine et al. (2011) has proposed that the *pilosae* lentil is derived from the Afghan germplasm. This is supported by genetic diversity studies by Ferguson et al. (1998) who established a close genetic affinity between the *pilosae* lentil and the Afghan germplasm. Erskine et al. (2011) has also suggested that the selection at intermediate

elevations for reduced sensitivity to photoperiod allowed for the spread into the short season environments of the Indo-Gangetic plain.

If the Indian landrace ILL 2601 is established to be vernalisation-insensitive, and *QTLA* is implicated in the control and regulation of the vernalisation response, it is reasonable to postulate that the incidence of the early haplotype in late flowering accessions could have facilitated the shift to spring-sowing which allowed for the introduction of lentils into Afghanistan and the surrounding region. This early haplotype of *QTLA* was likely maintained in the selection for photoperiod-insensitivity, afforded by *prr59c* as with ILL 2601, suited to the local agro-ecological environments of the Indo-Gangetic plain. Further work, including the study of the vernalisation response amongst these later flowering accessions carrying the early haplotype and a sequence analysis of *PRR59c* will provide an insight into the inconsistent late flowering phenotype of these accessions under non-inductive short days.

5.4.3 Molecular basis for QTLB

The flowering locus *QTLB* is established in Chapter 4 to confer ILL 2601 an early-flowering phenology. It is proposed that the locus is photoperiod-responsive, and controls both flowering time, and the interval between the appearance of a floral bud and the first developed flower.

In this chapter, it is proposed that *QTLB* is likely orthologous to the *Arabidopsis PRR9*. This chapter also suggests that an indel in *PRR59c*, which results in a frame-shift within the conserved CCT domain and a predicted premature stop codon, is likely responsible for the early-flowering phenology observed with ILL 2601 in SD. In lentils, a duplication of the *Arabidopsis* PRR9 orthologue is also proposed. Based on the phylogenetic analysis of the PRR5/9 clade in this chapter, this duplication is likely legume-specific.

In the current literature, the role of the *PRR59* gene family in legumes is not known. From limited work undertaken in *P. sativum*, it is proposed that the

PRR59a (formerly *PRR59*) is diurnally regulated, with elevated expression observed during the night (Liew et al., 2009a). It is not known if *PRR59c* is similarly regulated and if it functions as a component of the morning complex of the circadian clock. The *Arabidopsis* PRR5/9 clade function as components of the morning complex, with elevated expression observed during the night (Nakamichi et al., 2005).

The genes from the *PRR* gene family are generally associated with maintaining the rhythm of the circadian clock. Across both monocots (Beales et al., 2007; Cockram et al., 2007; Turner et al., 2005) and dicots (Nakamichi et al., 2012; Nakamichi et al., 2005; Nakamichi et al., 2007; Yamamoto et al., 2003), *PRR* mutants are associated with period lengthening, and a late-flowering phenotype. There is only one reported instance where a mutation for a *PRR* gene, *Ppd-1a* in *Triticum aestivum* (wheat), has been associated with photoperiod insensitivity and an early flowering phenology (Beales et al., 2007). The *ppd-d1a* mutant confers wheat a semi-dominant, early-flowering habit under SD (Beales et al., 2007; Boden et al., 2015), and has been extensively incorporated into existing breeding programs to develop wheat for short growing seasons where drought and heat stress are key considerations (Kato and Yokoyama, 1992). Mutations in the *T. aestivum Ppd-1a* gene however do not occur within the CCT domain (Beales et al., 2007; Boden et al., 2015).

5.4.4 Role of PRR59c in lentil

The lentil *prr59c* functions to afford photoperiod-insensitivity, and an early flowering phenology in ILL 2601. It is likely that the lentil *PRR59c* regulates the development of open flowers in response to changes in photoperiod, with SD conditions being deemed unfavourable for reproductive development, as observed with ILL 5588. It is likely that this mutation was selected for to allow the cultivation of lentil in the short-season, lower latitudes of the Indo-Gangetic plains.

More work is required to further verify the function of *PRR59c* in lentil, and its role in the adaptive evolution of early-flowering lentil.

5.4.5 Limitations of study

This chapter, while having significantly progressed the current understanding of the molecular basis for earliness in the Indian landrace ILL 2601, is limited in its scope to comprehensively describe the mode of action for each of the detailed mutation.

To further progress this study, a larger, more robust F_3 population will be required. The health and small size of both segregating populations severely limited the ability to draw more information regarding the studied loci, and record observations of NFI, DFD, and NFD.

A vernalisation experiment, aimed at understanding the vernalisation response of parental lines ILL 2601, ILL 5588, and their derivatives needs to be carried out to further validate *QTLA*, and its role in regulating flowering time independently of photoperiod. This potentially opens another perspective, alluded to in this chapter, in the control of flowering time in lentils.

The role of the lncRNA will also need to be further verified. Whilst more challenging, deriving a better appreciation for the mode of action of the lncRNA underlying *QTLA* affords the potential to better understand how flowering time is regulated independently of photoperiod.

The study with regards to *QTLB* needs to be progressed further. While it is likely that *PRR95c* is the candidate responsible for the observed phenotype, more work is required to reconcile the role of the *PRR* gene family in most monocots and dicots and the observations in lentil. Additionally, *MYB1* should be sequenced in both ILL 5588 and ILL 2601 to rule out its contribution to the observed early flowering phenotype.

Chapter 6 Characterising the late-flowering habit of cv. Indianhead

6.1 Introduction

The cv. Indianhead is a small-seeded (Muehlbauer et al., 1995), late-flowering (Vandenberg, A. 2012, pers. comm.), and strongly indeterminate (Hawtin et al., 1988) lentil cultivar. This cultivar was first selected for its use as a green manure or a plow-down crop to provide an alternative to summerfallow following cereal crops (Brandt, 1996; Slinkard, 1988). It is suggested that the value of nitrogen fixed by cv. Indianhead exceeds the cost of its production (Hawtin et al., 1988), and that as a cover crop it is more water-efficient than lucerne and sweet clover (Clark, 2008). Moreover, cv. Indianhead is suggested to be a low-cost green manure alternative when compared to other legume crops such as *Pisum sativum*, (Slinkard, 1988). The cultivar has been extensively incorporated into the production systems of Western Canada and the Northern Plains of the United States, where it is part of the cereal-legume rotation system (Allen et al., 2011; Zentner et al., 1996). In view of its purpose, an ideotype that favours vegetative development and increased dry matter production over seed production, conferred by a late-flowering phenology under inductive long-day conditions, would have best suited the requirements of the crop. It has been additionally proposed that cv. Indianhead as a green manure crop afforded the benefits of increased soil moisture content if the crop did not flower during development (Zentner et al., 1996).

More recently, cv. Indianhead has become increasingly incorporated into the food production system as a speciality crop. The small, rounded black seeds characteristic of cv. Indianhead are increasingly valued commercially, and is now marketed as a speciality class of lentil under the commercial name “Beluga Lentil” (McVicar et al., 2010; Muehlbauer et al., 2009).

The late-flowering phenology of cv. Indianhead has not been extensively studied nor described in the literature. Furthermore, the genetic basis for the late-flowering phenology of cv. Indianhead or other late flowering lentil accessions is not known, with the current literature (Sarker et al., 1999; Weller et al., 2012) and findings from chapters 3, 4, and 5 confined to the study of the early flowering habit.

Characterising the flowering phenology of cv. Indianhead allows for a better appreciation, and understanding of the control of flowering time, in particular late-flowering in lentil. Understanding and characterising the genetic control of the later-flowering phenology of cv. Indianhead, confers lentil breeders the capacity to exploit this economically valuable trait in their breeding programs, and provides an insight into genetic basis for the late-flowering phenology of lentil. Within the *Fabeae* tribe, the late-flowering phenology has been previously studied in the genus *Pisum*. Several late-flowering mutants have been genetically characterised in the *P. sativum* (Foucher et al., 2003; Hecht et al., 2007; Hecht et al., 2011), affording the opportunity for comparative genetics in the study of flowering time control in lentil.

This chapter aims to characterise the late-flowering phenology of cv. Indianhead. The chapter also investigates the genetic basis of the observed late-flowering phenology of cv. Indianhead. The chapter, while not seeking to determine the molecular basis of this phenotype, seeks to provide a foundation for future research with regards to lateness and the study of the genetic control of flowering time in lentil.

Experimental work and genetic analyses undertaken in this chapter were carried out in collaboration with the Department of Economic Development, Jobs, Transport and Resources, Victoria, Australia. The collaborative nature of the study and intellectual property of molecular marker information limited the full extent of analyses of phenotypic data carried out in this chapter. Linkage map illustration was not available for presentation at time of thesis submission.

6.2 Materials and methods

This section details specific materials and methods relevant to this chapter. General materials and methods are described in Chapter 2.

6.2.1 Plant materials and growth conditions

A total of 126 individuals from a recombinant inbred line (RIL) population derived from a cross between late flowering cv. Indianhead, and the photoperiod-sensitive cv. Northfield (ILL 5588) were evaluated under long day (LD) photoperiod conditions. Parental lines cv. Indianhead and ILL 5588 were evaluated under both LD and short day (SD) photoperiod conditions.

In this chapter, LD treated plants were exposed to 12-hour natural photoperiod supplemented with 4-hour fluorescent lighting, and SD treated plants were exposed to 12-hour natural photoperiod at the University of Tasmania phytotron.

RIL population was developed and provided by the Department of Economic Development, Jobs, Transport and Resources, Victoria, Australia.

Refer to section 5.2.1 and Appendix 5 for information pertaining to the University of Tasmanian in-house lentil accession collection.

6.2.2 Plant measurements

Refer to Chapter 2.

6.2.3 Genetic linkage map construction

A genetic linkage map was constructed by Kaur et al. (unpublished) from data obtained from single nucleotide polymorphism (SNP) and short repeat read (SSR) genotyping of 117 F_6 RIL derived from a cross between cv. Indianhead and ILL 5588. Map Manager QTXb19 (Manly et al., 2001) was utilized for the construction of the genetic linkage map. A total of 417 polymorphic markers were employed for map construction. Markers that displayed segregation

distortion ($p < 0.05$) were excluded from linkage map construction. The independence logarithm of odds (LOD) significance threshold was utilised in a manner of increasing stringency to assign statistically associated polymorphic markers into groups. A minimum LOD value of 6.0 and a recombination fraction (q) of 0.25 was used to assign markers to linkage groups. The marker order was verified using the *ripple* function. The Kosambi regression algorithm was employed to resolve the order of the polymorphic markers and the distances between markers within each group.

6.2.4 Quantitative trait loci (QTL) mapping

QTL mapping was carried out using QTL Cartographer v.2.5 (Wang et al., 2012). Both simple interval mapping (SIM) and composite interval mapping (CIM) were employed to determine loci attributable to traits analysed. A genome-wide LOD significance threshold for each trait was determined using 1000 permutations, and applied for QTL resolution.

6.3 Results

6.3.1 Characterisation of cv. Indianhead under different photoperiods

The flowering phenology of cv. Indianhead has not been previously reported nor described in the current literature. The late-flowering accession was evaluated in this study under both LD and SD photoperiod conditions. It was evaluated that cv. Indianhead is amongst the latest to flower under both photoperiod conditions when compared to 47 cultivated lentil accessions and one spp. *orientalis* accession, as illustrated in Figure 6-1A. Refer to section 5.2.1 and Appendix 5 for information pertaining to the University of Tasmanian in-house lentil accession collection.

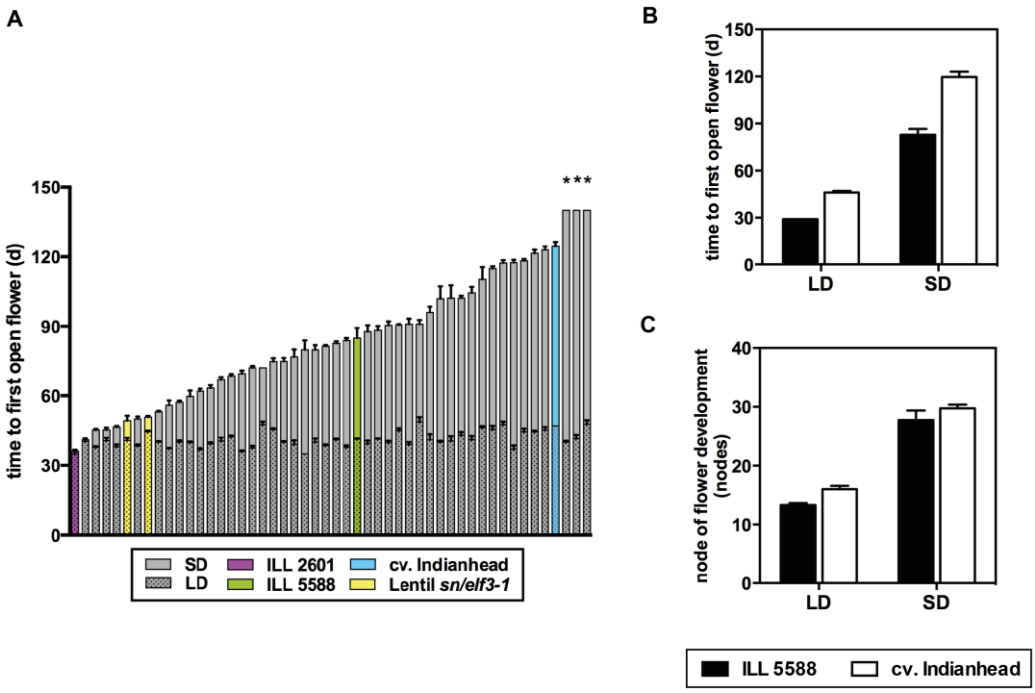


Figure 6-1 Phenotypic characterisation of cv. Indianhead under different photoperiods.
(A) DTF of University of Tasmania in-house lentil accessions under LD and SD. Asterisk (*) denotes accessions that did not flower after 140 days in SD. Accessions are arranged in order of SD flowering time. Accessions are illustrated as follows: cv. Indianhead (blue), ILL 5588 (green), ILL 2601 (pink), lentil *Sn* (yellow). (B) Flowering time, scored as days to first developed flower, in ILL 5588 and cv. Indianhead, under LD and SD conditions. (C) Node of flower initiation, denoting the physiological age at the point of transition to reproductive development, in ILL 5588 and cv. Indianhead, under LD and SD conditions. Plants received a 12-h photoperiod of natural daylight (SD) and a 16-h photoperiod of natural daylight (LD). Data are mean ±SE for *n*=3-4.

Cv. Indianhead flowered on average in 46.0 ± 1.00 days under inductive LD, significantly earlier ($p < 0.05$) than plants exposed to SD, which on average flowered in 120 ± 3.38 days (Figure 6-1B). Additionally, it was observed that under LD conditions, cv. Indianhead flowered at a significantly earlier ($p < 0.05$) developmental node (NFD = 16.0 ± 0.580 nodes) when compared to plants exposed to SD conditions (NFD = 29.8 ± 0.630 nodes) (Figure 6-1C). The reported flowering time response shows that while cv. Indianhead is later to flower when compared to ILL 5588 under both LD and SD, it remains photoperiod responsive. The photoperiod-sensitive ILL 5588 reported a DTF of 29.0 ± 0.00 days, and an average NFD of 13.3 ± 0.330 nodes under LD conditions, and a DTF of 82.9 ± 3.86 days, and an average NFD of 27.8 ± 1.65 nodes under SD conditions. In both LD and SD conditions, cv. Indianhead reported a later DTF ($p < 0.05$) than ILL 5588.

6.3.2 Flowering time segregation of cv. Indianhead X ILL 5588 RIL

A RIL population derived from a cross between the late-flowering cv. Indianhead and photoperiod-sensitive ILL 5588 was analysed in a controlled environment under a 16-hour photoperiod (Figure 6-2).

The phenotypic characterisation of the cv. Indianhead X ILL 5588 RIL population under LD conditions established an early class and a late class for DTF. The early class (mean DTF = 40.8, ±SD 11.5) was observed to flower significantly earlier ($p < 0.05$) than the late class as illustrated in Figure 6-2C.

However, within the early class, a large variation in DTF and NFD was reported, suggesting the role of multiple loci in the control of DTF within the early class (Figure 6-2B and Figure 6-2C). The late class reported a mean DTF of 86.7 days (±SD 5.15), with a 9.25 day interval between the latest flowering RIL from the early class and the earliest flowering RIL from the late class (Figure 6-2C). Individuals from the late class were additionally entirely transgressive, flowering on average 40.7 days later than cv. Indianhead in LD (Figure 6-2A).

The observation of transgressive early and late RIL suggests that there were multiple loci involved in the control of DTF in the cv. Indianhead x ILL 5588 RIL population (Figure 6-2A). It was also additionally observed that on average the early RIL transitioned to reproductive development at an earlier physiological age with the initiation of the first developed flower at a lower node (R^2_{adj} = 0.811), when compared to the late RIL (Figure 6-2B).

It was also observed that for all RIL individuals, the first floral structure developed into a fully developed flower. This is in contrast to observations reported in Chapter 4, where floral abortions and a delay to the first open flower were noted.

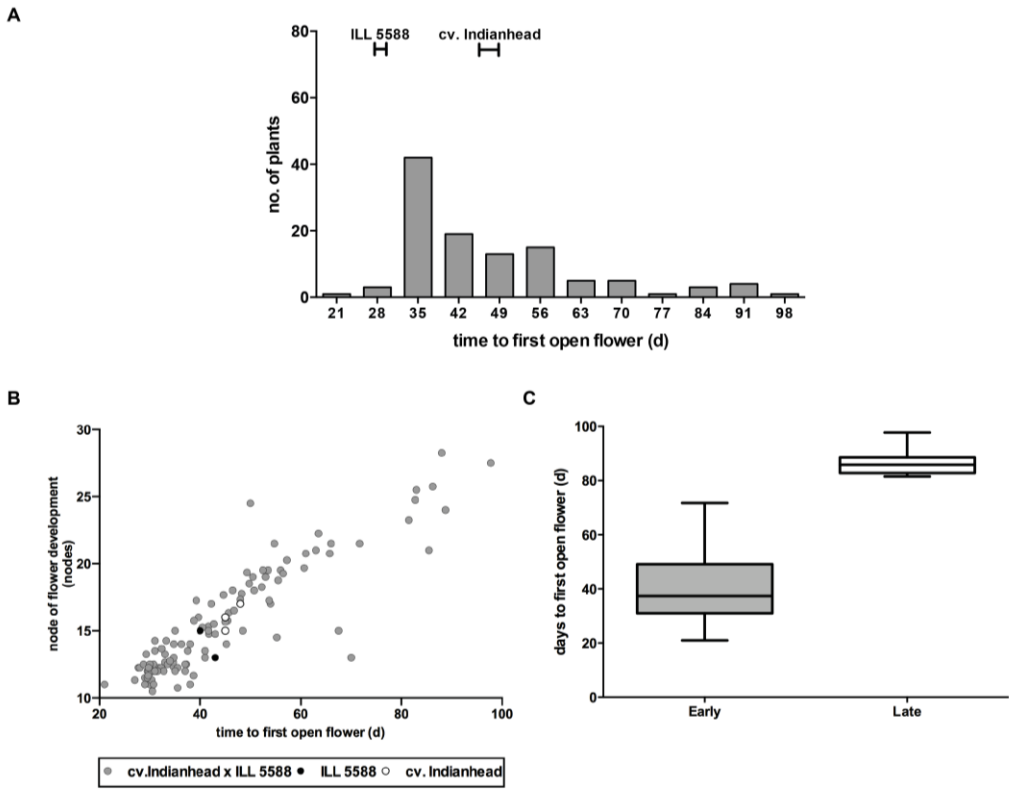


Figure 6-2 Segregation of cv. Indianhead x ILL 5588 RIL population for flowering time.
(A) Distribution of 126 RIL with respect to DTF under LD conditions. Data are for *n*=2-4. (B) Transition to reproductive development for 126 RIL illustrated in the context of DTF and NFD under LD conditions. Data are for *n*=2-4. (C) Mean DTF for Early and Late classes. Whiskers represent minimum and maximum DTF. Plants received a 16-h photoperiod of natural daylight.

6.3.3 QTL mapping for flowering time and candidate gene analysis

Two loci for DTF, and one locus for NFD were determined through QTL analysis in this experiment. For the purpose of this thesis, the DTF and NFD loci have been provisionally assigned *DTF3*, *DTF4*, and *NFD4* (Table 6-1).

		Simple Interval Mapping (SIM)						**Composite Interval Mapping (CIM)**					
Trait	**QTL**	**LOD threshold**	**Max LOD**	**Vp (%)**	**Linkage Group**	**Position**		**LOD threshold**	**Max LOD**	**Vp (%)**	**Linkage Group**	**Position**	
						Peak (cM)	**Range (cM)**					**Peak (cM)**	**Range (cM)**
DTF	*DTF3*	4.00	5.06	69.2	U	-	-	*not resolved*					
	DTF4	4.00	5.84	29.4	7	64.5	59.5/ 65.9	4.10	6.93	25.3	7	64.5	59.5/ 66.9
NFD	*NFD4*	3.60	6.95	52.8	7	62.5	59.5/ 65.9	3.80	10.3	40.8	7	62.5	60.5/ 64.9
	-		4.95	30.1	7	79.5	76.5/ 84.3		*not resolved*				

Table 6-1 Flowering time loci in cv. Indianhead x ILL 5588 RIL population.
QTL information for *days to flowering* (DTF) and *node of flower development* (NFD). LOD threshold determined using *permutation test*. Max LOD denotes the maximum LOD score for each trait at the determined QTL. Vp (%) denotes the estimated contribution of QTL to phenotypic variation. Peak (cM) is the linkage position of the peak marker, and range (cM) denotes the 1-LOD interval of the QTL. The syntenic relationship of the linkage group nomenclature is detailed in Appendix 9.

DTF3 contributes to an estimated 69.2% (SIM) of observed variation for DTF and occurs in the region syntenic to *M. truncatula* chromosome 8, while *DTF4* is estimated to contribute to 29.4% (SIM) and 25.3% (CIM) of the observed variation for DTF and occurs in the region syntenic to *M. truncatula* chromosome 3 in the cv. Indianhead x ILL 5588 RIL population under LD (Table 6-1). Additionally, *NFD4* was analysed to occur within the 2-LOD interval of *DTF4*, suggesting that it is likely that both loci are co-located. No NFD locus was determined for the peak marker associated with *DTF3*.

DTF3 was determined to be associated to the PBA_LC_083_IH marker when the SIM procedure was employed during the first round of QTL mapping. The PBA_LC_083_IH marker is an unmapped marker (U), dominant for cv.

Indianhead, which was not incorporated into co-dominant framework map for the cv. Indianhead x ILL 5588 RIL population (Kaur et al., unpublished). The secondary CIM procedure excludes unmapped markers during QTL analysis, and hence *DTF3* was not determined during the secondary QTL mapping.

The PBA_LC_083_IH marker was subsequently mapped on the linkage group corresponding to *M. truncatula* chromosome 8 in a draft framework map that included both dominant and co-dominant markers (Kaur et al., unpublished).

Through comparative mapping, several candidate genes, associated with flowering time in other systems, were identified for each of the identified DTF and NFD loci (Table 6-2). Each of these candidates is predicted to occur within the 2-LOD confidence interval of the QTL.

Locus	Gene symbol	Medicago locus (Mt4.0)	Gene description	Additional information
DTF3	*Jmj14*	Medtr8g089260	lysine-specific demethylase JMJ14-like protein, putative	Involved in flowering time regulation in *Arabidopsis*. Involved in posttrascriptional gene regulation (Lu et al., 2010).
	AGa	Medtr8g087860	MADS-box transcription factor	*Closest hit to AGAMOUS (AG). AG* is implicated in floral development in *Arabidopsis* (Bowman et al., 1991).
	Myb-like	Medtr8g086410	MYB-like transcription factor family protein	*P. sativum MYB1 demonstrates strong diurnal expression rhythm in LD and SD* (Liew et al., 2009a). *P. sativum MYB1 is a Arabidopsis CCA1 orthologue.*
DTF4*/ *NFD4	*TEJ*	Medtr3g029520	poly(ADP-ribose) glycohydrolase	*tej* mutants increase period length in *Arabidopsis*. (Panda et al., 2002)
	bHLH122-like	Medtr3g027650	bHLH-like transcription factor family protein	*AtbHLH* functions as a *CONSTANS* (*CO*) transcriptional activator and regulates flowering time (Ito et al., 2012).
	AGL62	Medtr3g030780	MADS-box transcription factor family protein	*Closest hit to AGAMOUS-like 62 (AGL62). AG is implicated in floral development in Arabidopsis* (Bowman et al., 1991).

Table 6-2 Candidate genes for *DTF3*, *DTF4*, and *NFD4*
Candidate genes positioned within the 2-LOD vicinity of peak position are listed and described.

6.3.4 Other quantitative traits

The cv. Indianhead x ILL 5588 RIL population was also evaluated for several other quantitative traits under LD. These included *time to emergence* (DTE), *number of pods per node* (PPN), *number of branches to DTF* (BTF), and internode length.

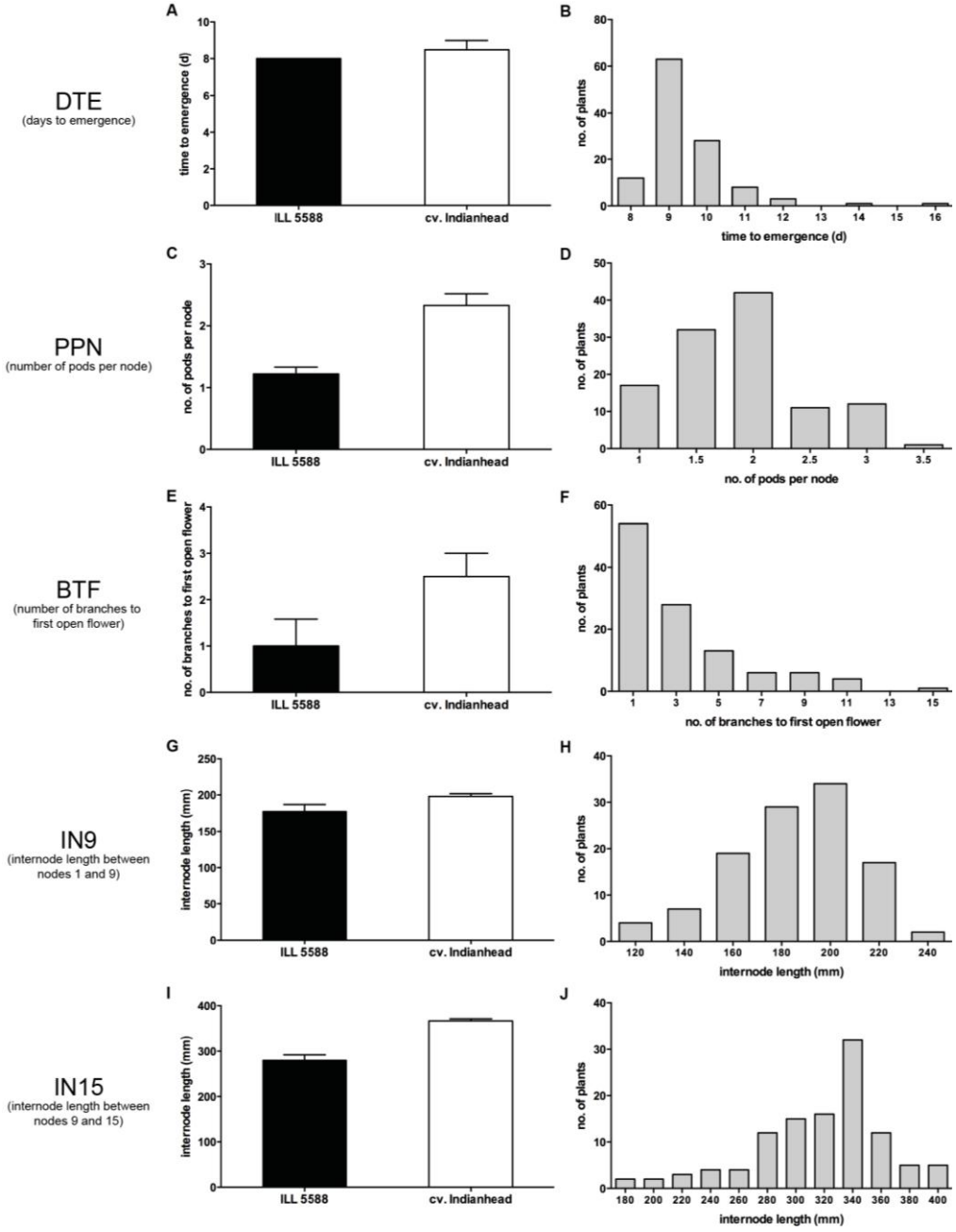


Figure 6-3 Segregation of cv. Indianhead x ILL 5588 RIL population for other traits.
(A) Days to emergence in ILL 5588 and cv. Indianhead. (B) Distribution of days to emergence. (C) Average number of pods per node observed from R1 to R3. (D) Distribution of the average number of pods per node observed from R1 to R3. (E) Total number of branches observed to first developed flower. (F) Distribution of the total number of branches observed to first developed flower. (G) Internode length (mm) between nodes 1 and 9. (H) Distribution of the

observed internode length (mm) between nodes 1 and 9. (I) Internode length (mm) between nodes 9 and 15. (J) Distribution of the observed internode length (mm) between nodes 9 and 15 in cv. Indianhead X ILL 5588 RIL population. Plants received a 16-h photoperiod of natural daylight. Data are mean ±SE for *n*=3-4. 126 RIL were evaluated in study.

Amongst the quantitative traits analysed, it was reported that the cv. Indianhead was observed to produce more pods per node ($p < 0.05$) (Figure 6-3C), more branches to first developed flower ($p < 0.05$) (Figure 6-3E), and had a longer internode interval between nodes 9 and 15 ($p < 0.05$) (Figure 6-3I), when compared to ILL 5588 (Figure 6-3). There was no significant difference observed in the time to emergence from sowing ($p = 0.356$) (Figure 6-3A), and internode interval between nodes 1 and 9 between both genotypes ($p = 0.384$) (Figure 6-3G).

6.3.5 QTL mapping for other quantitative traits

Loci for several quantitative traits including days to emergence (DTE), pods per node (PPN), branches to flowering (BTF), internode interval between nodes 9 and 15 (IN15) were determined during QTL analysis. Interestingly, it was also determined that the single locus attributable to BTF is co-located (2-LOD interval) with both *DTF4* and *NFD4*. Loci determined for above traits are summarised in Table 6-3.

Trait	QTL	Simple Interval Mapping (SIM)						Composite Interval Mapping (CIM)					
		LOD threshold	Max LOD	Vp (%)	Linkage Group	Position Peak (cM)	Position Range (cM)	LOD threshold	Max LOD	Vp (%)	Linkage Group	Position Peak (cM)	Position Range (cM)
Days to emergence (DTE)	-	2.60	3.52	17.9	3	65.6	58.6/65.8	2.60	*not resolved*				
	not assigned		3.86	19.8	3	80.3	79.6/84.2		4.25	18.4	3	80.3	78.6/83.3
Pods per node (PPN)	not assigned	3.10	5.83	29.4	4	8.00	4.6/14.7	3.20	6.76	27.1	4	8.00	4.60/14.7
	-		4.16	22.5	4	22.1	19.2/25.0		*not resolved*				
	not assigned	*not resolved*						3.20	4.27	12.9	5	50.4	44.6/54.8
Branches to flowering (BTF)	not assigned	2.90	*not resolved*					4.00	5.44	19.6	7	63.5	57.5/65.9
Internode interval between nodes 1 and 9 (IN9)	-	3.10	*not resolved*					3.10	*not resolved*				
Internode interval between nodes 9 and 15 (IN15)	not assigned	3.10	*not resolved*					3.00	3.19	12.1	4	159	96.9/94.6
	not assigned							3.00	3.57	13.4	7	55.6	49.6/63.5

Table 6-3 Summary of QTL determined for quantitative traits in cv. Indianhead x ILL 5588 RIL population.
LOD threshold determined using *permutation test*. Max LOD denotes the maximum LOD score for each trait. Vp (%) denotes the estimated contribution of QTL to phenotypic variation. Peak (cM) is the linkage position of the peak marker, and range (cM) denotes the 1-LOD interval of the QTL. The syntenic relationship of the linkage group nomenclature is detailed in Appendix 9.

6.4 Discussion

The flowering phenology of cv. Indianhead has not been documented nor described in the current literature. It has been suggested that the largely indeterminate (Hawtin et al., 1988) cultivar, first developed as a low-cost green manure (Slinkard, 1988), and an alternative to summerfallow (Brandt, 1996), is extremely late to flower. In this chapter, the late-flowering phenology of the cultivar is determined to be photoperiod responsive, and cv. Indianhead is described to represent amongst the latest flowering accession within a collection that encapsulates the broad range of flowering times observed in cultivated lentil.

6.4.1 Genetic basis for late-flowering phenology

In this chapter, it is proposed that three loci, namely *DTF3*, *DTF4* and *NFD4*, collectively contribute to the observed late-flowering phenology. *DTF3* is reported to account for the largest proportion of the observed variation in the segregating RIL population, and is proposed to occur in the region syntenic to *M. truncatula* chromosome 8. No loci contributing to NFD variation was observed to co-locate with *DTF3. DTF4*, which accounts for the residual variation for DTF is co-located with *NFD4*, and is positioned on the region syntenic to *M. truncatula* chromosome 3. Both loci are also suggested not to occur in the region of the three currently established lentil flowering time loci, *Sn*/*ELF3*, *DTF1*/*QTLA*, and *DTF2*/*QTLB*.

6.4.2 Candidate genes analysis

In *P. sativum*, there has been significant progress in the understanding of the genetic and molecular basis for the late flowering-phenology. The late flowering phenology has been previously attributed to genes at the *P. sativum GIGAS*, *LF*, and *LATE1* loci (Foucher et al., 2003; Hecht et al., 2007; Hecht et al., 2011; Weller et al., 2012). However, based on QTL mapping and the resolution of the two loci in cv. Indianhead, and the established macrosyntenic relationship between both lentil and *P. sativum* with *M. truncatula*, it is unlikely that an

orthologous gene at either of these loci are responsible for the observed phenotype.

The established synteny between lentil and *M. truncatula* provides a secondary basis for the nomination of candidate genes for both *DTF3* and *DTF4*/*NFD4*. Several genes associated with flower development and flowering time in other systems were analysed to occur within the vicinity of the *DTF3* peak marker. Amongst the nominated candidates, *JMJ14*, a member of the Jumonji C (JmjC) family is of interest. *JMJ14* is a H3K4 demethylase that has been previously reported to regulate flowering time in *Arabidopsis* by increasing the relative expression of downstream genes associated with flowering including *FT*, *SOC1*, *LFY*, and *AP1* (Lu et al., 2010). The relationship between the expression of lentil *FT* orthologues and the time to flower were validated in Chapter 5, and the proposition that the candidate *JMJ14* is involved in the regulation of *FT* is noteworthy. Furthermore, it is also reported that *JMJ14* function to regulate flowering time through posttranscriptional gene silencing, and that mutants prevent the silencing of regulated genes (Le Masson et al., 2012). Lu et al. (2010) has also proposed that while involved in flowering time regulation, *jmj14* mutants remained photoperiod sensitive, similar to cv. Indianhead.

The second locus *DTF4*, which is co-located with *NFD4*, occurs in the region that corresponds to *M. truncatula* chromosome 3. Several genes associated with flower development or flowering time are positioned within the 2-LOD confidence interval of the locus. The *Arabidopsis TEJ* gene, a poly(ADP-ribose) glycohydrolase, associated with period lengthening is one of the identified candidates (Panda et al., 2002). The *TEJ* gene is involved in the regulation of the circadian clock and its mutants are early to flower (Panda et al., 2002). The function of the *TEJ* orthologue in legume systems is not known. Another candidate identified is a basic helix-loop-helix-type transcription factor (bHLH). Several flowering bHLH transcriptional activators have been idenitifed previously in *Arabidopsis*, and it is suggested that bHLHs are involved in the regulation of *CO* (Ito et al., 2012). While *CO* and its homologues has been

established in *M. truncatula* to not be important in the regulation of flowering time (Wong et al., 2014), the role of these transcriptional activators in the control of flowering time have not been studied in lentil or more broadly in the legume system.

This chapter does not progress beyond the nomination of candidate genes, associated with flowering time in other systems. More work is required to further probe the molecular basis for these two loci.

6.4.3 Future directions for study

This chapter determined two new loci involved in the control of a late-flowering phenology and significantly progresses the current understanding of the genetic control of flowering time in lentil.

The study of RIL families segregating for each of the identified loci will provide the basis for a candidate-gene approach to resolving the molecular basis for both *DTF3* and *DTF4*/*NFD4*.

The study of these loci under SD photoperiod conditions will further contribute to the findings of this chapter, and expand the understanding of the role of these loci in the control of flowering time.

A survey of a representative lentil collection for the prevalence of these loci will further inform the origin and contribution of the late alleles to the variation in flowering time for cultivated lentil.

Chapter 7 General discussion

This thesis investigated the genetic and molecular basis for flowering time control in lentil. Findings from this thesis progresses the limited understanding of this process presented in the literature, and expands the current premise that in lentil flowering time is controlled by a polygenic system shaped by interactions between the major lentil *Sn* locus and several minor loci. The thesis also presented new perspectives on the regulation of flowering time in lentil, and afforded an insight on the genetic basis for flowering time adaptation.

7.1 Update on the genetic control of flowering time in lentil

Flowering time is often described in the context of *days to flowering* (DTF) and the *node of flower development* (NFD). This thesis determined a moderate (Chapter 3) to strong (Chapter 4 and 6) positive correlation between both DTF and NFD in lentil. Chapter 4 also determined that NFD in lentil is a complex trait, and postulates that two independent traits; *node of floral initiation* (NFI) and *delay to flower development* (DFD) contribute to the observed variation for NFD (refer to Figure 7-1A for illustration).

7.1.1 Photoperiod-independent regulation

Findings from this thesis propose that in lentil each of the defined node-based traits for flowering are a function of individual pathways that synergistically regulate the shift to an optimal flowering phenology. In Chapter 4, a novel locus designated *QTLA*, is described to control DTF (Figure 4-6) and the developmental node for NFI (Figure 4-12). The chapter also determined that NFI is regulated independently of photoperiod.

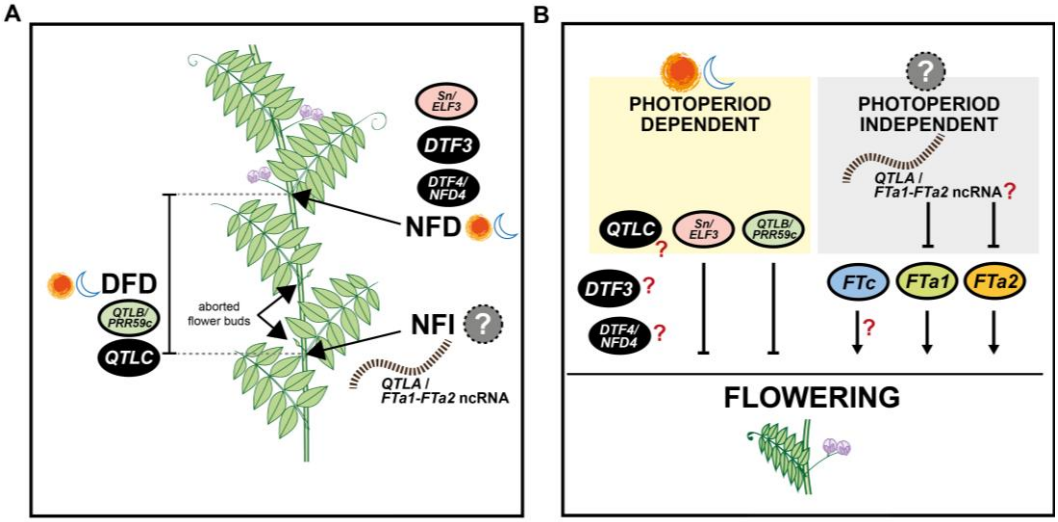


Figure 7-1 Proposed model for the genetic control of flowering time in lentil.
(A) Traits contributing to flowering time variation. Loci implicated in the control of flowering time are illustrated. NFD refers to the *node of flower development*, NFI refers to the *node of flower initiation*, DFD refers to the *node interval between NFI and NFD*. (B) The mode of action of loci determined in this thesis is illustrated. Arrows indicate loci promoting flowering, while bars loci inhibiting flowering. In both figures, 'black-white' loci indicate unknown molecular basis, the 'sun and moon' symbols indicate photoperiodic basis, and red '?' indicate that more work is required.

Chapter 5 proposes that expressed non-coding RNA (ncRNA) form the molecular basis for *QTLA* (Figure 5-6), and that it acts through the regulation of *FTa1* and *FTa2* (Figure 5-3 and Figure 7-1B). It is not known how this is achieved, but precedence in the mice system point to the likely epigenetic regulation of the ncRNA that facilitate the promotion and repression of neighbouring genes.

The proposed photoperiod-independent pathway, and its control of the first floral structure, draws interesting parallels to observations of a photoperiod-insensitive pre-inductive phase described by Roberts et al. (1986) for lentils. Roberts et al. (1986) had inferred that this phase is the juvenile phase or basic vegetative phase, and his findings point to a variation (time) for his phase across evaluated accessions.

The elevated expression of both *FTa1* and *FTa2* during early development of early flowering ILL 2601 (Figure 5-3) furthermore mirrors observations by Jaudal et al. (2013) of *M. truncatula* vernalisation mutants. Jaudal et al. (2013) additionally determined that retroelement insertions 3' of *FTa1* can eliminate

vernalisation requirements and confer a dominant early-flowering phenology in *M. truncatula*. This resembles the dominant flowering habit conferred by the ILL 2601 haplotype for the *FTa1-FTa2* intergenic region (Figure 5-7). While this thesis does not explore vernalisation, Summerfield et al. (1985) and Roberts et al. (1986) have determined that in lentil vernalisation can reduce the nominal base photoperiod required for flower induction, and shorten the critical photoperiod required for flowering.

It is of interest to the study of flowering time in lentil to validate the role of the novel *QTLA* in the vernalisation response. Findings from Chapters 4 and 5 present opportunities for future work in the area.

7.1.2 Photoperiod-dependent regulation

The existing literature on the photoperiodic basis for flowering (Roberts et al., 1988; Roberts et al., 1986; Summerfield et al., 1985), and flowering time variation (Erskine et al., 1990a; Erskine et al., 1994), imply the presence of a photoperiod-dependent pathway for flowering time in lentil. Inference from previous work on the lentil *Sn* (Sarker et al., 1999) and cv. Precoz (Roberts et al., 1986) suggests a photoperiodic basis for the previously characterised flowering time locus. Chapter 3 determined that the lentil *Sn* locus is orthologous to the *Arabidopsis ELF3* circadian clock gene (Figure 3-5), and that the recessive *elf3-1* alleles derived from cv. Precoz afford an early flowering phenology by conferring photoperiod-insensitivity (Figure 3-4).

Chapter 4 defines two novel lentil flowering time loci, designated *QTLB* and *QTLC*, implicated in the photoperiod-dependent pathway (Figure 4-1 and Figure 4-6). Chapter 5 proposes that the newly defined *QTLB* flowering time locus is a legume-specific *PRR* paralogue and that mutant *prr59c* alleles from ILL 2601 similarly afford an early flowering phenology by conferring photoperiod-insensitivity. While it can be inferred from other systems that the legume-specific *PRR* paralogue acts to regulate the circadian rhythm, its mode of action remains unclear.

7.1.3 Interplay between flowering pathways

The development of the first open flower as illustrated in Figure 7-1A, begins with the initiation of a floral structure at NFI. However, findings from Chapter 4 and anecdotal observations in Chapter 3 show that this floral structure is unlikely to develop into an open flower in photoperiod-responsive plants (ILL 5588) under non-inductive short days.

A delay (DFD) characterised by repeated floral abortions is suggested to occur under non-inductive short days (Figure 7-1A). Plants with photoperiod defects (ILL 6005 and ILL 2601) do not demonstrate floral abortions under short days; instead develop an open flower at the first reproductive node. Similarly, photoperiod-responsive plants (ILL 5588 and cv. Indianhead) are observed to develop an open flower at the first reproductive node under inductive long days. The interplay between the both pathways manifests through DFD.

Observations from this thesis further point to the requirement for inductive photoperiod for flowering, consistent with findings by Roberts et al. (1986) and Summerfield et al. (1985) that determined that while the critical base photoperiod can be reduced, presumably by *QTLA*, the requirement for photoperiod cannot be eliminated. The photoperiod-dependent basis for DFD also mirrors findings by Jaudal et al. (2013), that in plants with retroelement insertions 3' of *FTa1*, there still is a requirement for inductive photoperiod.

It is not known how these pathways interact or synergistically act to regulate flowering time. Chapter 5 introduces the role of *Arabidopsis FT* orthologues in lentil flowering time regulation. It had been reviewed by Weller et al. (2015) that in *P. sativum* and *G. max*, these floral promoters act as floral integrators. It is suggested that both photoperiod (Hecht et al., 2011) and vernalisation (Laurie et al., 2011) are involved in the regulation of legume-specific *FTs*. Findings from Chapter 5 suggest that *QTLA* acts on *FTa1* and *FTa2*. *FTc* however is likely to be further downstream of the flowering pathway, as its expression mirrors the appearance of flower buds in lentil.

Separately, Chapter 6 of this thesis introduces two new loci involved in conferring a late-flowering phenology under inductive long day photoperiods. It is unclear from findings in this thesis how these loci interact or act in relation to the defined loci and described flowering pathways.

7.2 Genetic control of flowering time adaptation

In the evaluation of a collection of cultivated lentil of diverse geographic origins, this thesis reports a wide variation for flowering time in long and short day photoperiods (Figure 5-7D). This is consistent with observations by Erskine et al. (1989) and Erskine et al. (1990a). This thesis explored the molecular basis for this wide variation by surveying for the prevalence of the early-flowering alleles of the lentil *Sn* (Figure 3-7) and *QTLA* (Figure 5-7). This thesis presents the first attempt at deciphering the molecular basis for flowering time adaptation in lentil.

Chapter 3 of this study examined the prevalence of the mutant *elf3-1* allele in a Lentil Association Mapping panel (Figure 3-7), and proposes that the lentil *Sn/ELF3* is unlikely responsible for the early flowering phenotypes of both the *pilosae* and *aethiopicae* lentil.

Chapter 5 surveyed for the prevalence of the early haplotype for *QTLA* in a collection of cultivated lentil of diverse geographical origins (Figure 5-7) and predicts that the early haplotype for *QTLA* facilitated the introduction of cultivated lentil to the higher latitudes of Afghanistan and its surrounding region, where spring-sowing is practiced. This was predicted based on the likely role of *QTLA* in regulating the vernalisation response. Chapter 5 additionally inferred that the selection for loci conferring photoperiod-insensitivity, similar to *prr59c*/*qtlb*, was responsible for the eventual dissemination of cultivated lentil into the Indo-Gangetic plain.

7.3 Future work

The proposed model for flowering time presented in this thesis is the first reported attempt at integrating existing literature with more recent findings (from this study) to describe the genetic control of flowering time in lentil. This model presents numerous opportunities for future work.

Deciphering the control and regulation of *QTLA* is the most significant opportunity presented through findings from this thesis. The mode of action of *QTLA* still remains unresolved, and its role in the regulation of *FTa1* and *FTa2* in the context of vernalisation needs to be further studied.

The function of the legume-specific *PRR* paralogue that forms the molecular basis for *QTLB* also needs to be further studied. While this thesis describes the likely polymorphism in *prr59c* responsible for the mutant phenotype, more work is required to reconcile this observation with the role of the PRR5/9 family in most monocots and dicots. Diurnal expression studies and complementation studies can be explored to achieve this.

Chapter 6 of this thesis introduces two new loci involved in the genetic control of the late-flowering phenology of cv. Indianhead in inductive photoperiods. Presented candidates in the study provide a starting point for a candidate-gene approach to determining the molecular identity of these loci. More work is also required to reconcile the role of these loci in the proposed model for flowering time in lentil.

Separate to future work pertaining to flowering time, this thesis also presented an updated iteration of the syntenic relationship between the seven linkage groups of lentils and the eight chromosomes of *M. truncatula*. However, findings from this study only present a limited understanding of the relationship between *M. truncatula* chromosome 6 and lentil linkage group 2. More work is required to resolve the extent of translocation and the order of genes for this region.

References

Abbo, S., and Ladizinsky, G. (1991). Anatomical aspects of hybrid embryo abortion in the genus *Lens* L. Botanical gazette, 316-320.

Abbo, S., Saranga, Y., and Peleg, Z. (2009). Reconsidering domestication of legumes versus cereals in the ancient Near East. The Quarterly Review of Biology *84*, 29-50.

Ahmad, M., McNeil, D.L., Fautrier, A.G., Armstrong, K.F., and Paterson, A.M. (1996). Genetic relationships in *Lens* species and parentage determination of their interspecific hybrids using RAPD markers. Theor Appl Genet *92*, 1091-1098.

Allen, B.L., Pikul, J.L., Waddell, J.T., and Cochran, V.L. (2011). Long-Term Lentil Green-Manure Replacement for Fallow in the Semiarid Northern Great Plains. Agronomy Journal *103*, 1292-1298.

Alo, F., Furman, B.J., Akhunov, E., Dvorak, J., and Gepts, P. (2011). Leveraging genomic resources of model species for the assessment of diversity and phylogeny in wild and domesticated lentil. Journal of Heredity *102*, 315-329.

Amasino, R. (2010). Seasonal and developmental timing of flowering. The Plant Journal *61*, 1001-1013.

Barulina, H. (1930). Lentils of the USSR and other countries. Bulletin of Applied Botany, Genetics & Plant Breeding, Leningrad (Supplement) *40*, 265-304.

Beales, J., Turner, A., Griffiths, S., Snape, J., and Laurie, D. (2007). A *Pseudo-Response Regulator* is misexpressed in the photoperiod insensitive *Ppd-D1a* mutant of wheat (Triticum aestivum L.). Theor Appl Genet *115*, 721-733.

Berghoff, E.G., Clark, M.F., Chen, S., Cajigas, I., Leib, D.E., and Kohtz, J.D. (2013). *Evf2* (*Dlx6as*) lncRNA regulates ultraconserved enhancer methylation and the differential transcriptional control of adjacent genes. Development (Cambridge, England) *140*, 4407-4416.

Berry, G., and Aitken, Y. (1979). Effect of photoperiod and temperature on flowering in pea (*Pisum sativum* L.). Australian Journal of Plant Physiology *6*, 573-587.

Bett, K., and Cook, D. (2015). Lentil genome pre-release v0.8. In The Lentil Genome Sequencing (LenGen) Project.

Bett, K., Ramsay, L., Sharpe, A., Cook, D., Penmetsa, R., Stonehouse, R., Wong, M., Chan, C., Vandenberg, A., VanDeynze, A*., et al.* (2014). Lentil genome sequencing: establishing a comprehensive platform for molecular breeding. In International Food Legumes Research Conference VI and

International Conference on Legume Genetics and Genomics VII (Saskatoon, Canada).

Boden, S.A., Cavanagh, C., Cullis, B.R., Ramm, K., Greenwood, J., Jean Finnegan, E., Trevaskis, B., and Swain, S.M. (2015). *Ppd-1* is a key regulator of inflorescence architecture and paired spikelet development in wheat. Nature Plants *1*, 14016.

Bond, A.M., VanGompel, M.J.W., Sametsky, E.A., Clark, M.F., Savage, J.C., Disterhoft, J.F., and Kohtz, J.D. (2009). Balanced gene regulation by an embryonic brain ncRNA is critical for adult hippocampal GABA circuitry. Nat Neurosci *12*, 1020-1027.

Bowman, J.L., Drews, G.N., and Meyerowitz, E.M. (1991). Expression of the Arabidopsis floral homeotic gene *AGAMOUS* is restricted to specific cell types late in flower development. The Plant Cell *3*, 749-758.

Brandt, S.A. (1996). Alternatives to summerfallow and subsequent wheat and barley yield on a Dark Brown soil. Canadian Journal of Plant Science *76*, 223-228.

Clark, A. (2008). Managing Cover Crops Profitably (3rd Ed.) (Diane Publishing Company).

Cockram, J., Jones, H., Leigh, F.J., O'Sullivan, D., Powell, W., Laurie, D.A., and Greenland, A.J. (2007). Control of flowering time in temperate cereals: genes, domestication, and sustainable productivity. Journal of Experimental Botany *58*, 1231-1244.

Colledge, S., Conolly, J., and Shennan, S. (2005). The evolution of Neolithic farming from SW Asian origins to NW European limits. European Journal of Archaeology *8*, 137-156.

Covell, S., Ellis, R.H., Roberts, E.H., and Summerfield, R.J. (1986). The influence of temperature on seed germination rate in grain legumes: I. a comparison of chickpea, lentil, soyabean and cowpea at constant temperatures. Journal of Experimental Botany *37*, 705-715.

Cubero, J.I. (1984). Taxonomy, distribution and evolution of the lentil and its wild relatives. In Genetic resources and their exploitation: chickpeas, faba beans and lentils, J. Witcombe, and W. Erskine, eds. (Springer Netherlands), pp. 187-203.

Cubero, J.I., Pérez de la Vega, M., and Fratini, R. (2009). Origin, phylogeny, domestication and spread. In The lentil: botany, production and uses, W. Erskine, F.J. Muehlbauer, A. Sarker, and B. Sharma, eds. (Wallingford, UK, CABI), pp. 13-33.

de Candolle, A. (1882). Origin of cultivated plants (New York, D. Appleton and company).

Dias, P., Brunel-Muguet, S., Durr, C., Huguet, T., Demilly, D., Wagner, M.-H., and Teulat-Merah, B. (2011). QTL analysis of seed germination and pre-emergence growth at extreme temperatures in *Medicago truncatula*. Theor Appl Genet *122*, 429-444.

Doebley, J.F., Gaut, B.S., and Smith, B.D. (2006). The molecular genetics of crop domestication. Cell *127*, 1309-1321.

Durán, Y., Fratini, R., García, P., and Pérez de la Vega, M. (2004). An intersubspecific genetic map of *Lens*. TAG Theoretical and Applied Genetics *108*, 1265-1273.

Ellwood, S., Phan, H., Jordan, M., Hane, J., Torres, A., Avila, C., Cruz-Izquierdo, S., and Oliver, R. (2008). Construction of a comparative genetic map in faba bean (*Vicia faba* L.); conservation of genome structure with *Lens culinaris*. BMC Genomics *9*, 380.

Emami, M.K., and Sharma, B. (1999). Linkage between three morphological markers in lentil. Plant Breeding *118*, 579-581.

Erskine, W., Adham, Y., and Holly, L. (1989). Geographic distribution of variation in quantitative traits in a world lentil collection. Euphytica *43*, 97-103.

Erskine, W., Bayaa, B., and Saxena, M. (1996). Registration of ILL 5588 lentil germplasm resistant to Vascular Wilt and Ascochyta Blight. Crop Science *36*, 1080-1080.

Erskine, W., Chandra, S., Chaudhry, M., Malik, I.A., Sarker, A., Sharma, B., Tufail, M., and Tyagi, M.C. (1998). A bottleneck in lentil: widening its genetic base in South Asia. Euphytica *101*, 207-211.

Erskine, W., Ellis, R.H., Summerfield, R.J., Roberts, E.H., and Hussain, A. (1990a). Characterization of responses to temperature and photoperiod for time to flowering in a world lentil collection. Theor Appl Genet *80*, 193-199.

Erskine, W., Hussain, A., Tahir, M., Bahksh, A., Ellis, R.H., Summerfield, R.J., and Roberts, E.H. (1994). Field evaluation of a model of photothermal flowering responses in a world lentil collection. TAG Theoretical and Applied Genetics *88*, 423-428.

Erskine, W., Muehlbauer, F.J., and Short, R.W. (1990b). Stages of development in lentil. Experimental Agriculture *26*, 297-302.

Erskine, W., Sarker, A., and Ashraf, M. (2011). Reconstructing an ancient bottleneck of the movement of the lentil (*Lens culinaris* ssp. *culinaris*) into South Asia. Genetic Resources and Crop Evolution *58*, 373-381.

Eujayl, I., Baum, M., Erskine, W., Powell, W., and Pehu, E. (1998). A genetic linkage map of lentil (*Lens* sp.) based on RAPD and AFLP markers using recombinant inbred lines. Theor Appl Genet *97*, 83-89.

Faure, S., Turner, A.S., Gruszka, D., Christodoulou, V., Davis, S.J., von Korff, M., and Laurie, D.A. (2012). Mutation at the circadian clock gene *EARLY MATURITY 8* adapts domesticated barley (*Hordeum vulgare*) to short growing seasons. Proceedings of the National Academy of Sciences.

Fedoruk, M.J. (2013). Linkage and association mapping of seed size and shape in lentil. In Department of Plant Sciences (Saskatoon, University of Saskatchewan).

Feng, J., Bi, C., Clark, B.S., Mady, R., Shah, P., and Kohtz, J.D. (2006). The *Evf-2* noncoding RNA is transcribed from the *Dlx-5/6* ultraconserved region and functions as a *Dlx-2* transcriptional coactivator. Genes & Development *20*, 1470-1484.

Ferguson, M., Robertson, L., Ford-Lloyd, B., Newbury, H.J., and Maxted, N. (1998). Contrasting genetic variation amongst lentil landraces from different geographical origins. Euphytica *102*, 265-273.

Ford, R., Pang, E.C.K., and Taylor, P.W.J. (1997). Diversity analysis and species identification in *Lens* using PCR generated markers. Euphytica *96*, 247-255.

Foucher, F., Morin, J., Courtiade, J., Cadioux, S., Ellis, N., Banfield, M.J., and Rameau, C. (2003). *DETERMINATE* and *LATE FLOWERING* Are Two *TERMINAL FLOWER1/CENTRORADIALIS* Homologs That Control Two Distinct Phases of Flowering Initiation and Development in Pea. The Plant Cell *15*, 2742-2754.

Fratini, R., and Ruiz, M. (2006). Interspecific hybridization in the genus *Lens* applying in vitro embryo rescue. Euphytica *150*, 271-280.

Fuller, D.Q. (2007). Non-human genetics, agricultural origins and historical linguistics in South Asia. In The Evolution and History of Human Populations in South Asia, M.D. Petraglia, and B. Allchin, eds. (Springer Netherlands), pp. 393-443.

Fuller, D.Q., Asouti, E., and Purugganan, M.D. (2011a). Cultivation as slow evolutionary entanglement: comparative data on rate and sequence of domestication. Vegetation History and Archaeobotany, 1-15.

Fuller, D.Q., Boivin, N., Hoogervorst, T., and Allaby, R. (2011b). Across the Indian Ocean: the prehistoric movement of plants and animals. Antiquity *85*, 544-558.

Galasso (2003). Distribution of highly repeated DNA sequences in species of the genus *Lens* Miller. Genome *46*, 1118.

Gupta, D., Taylor, P.W.J., Inder, P., Phan, H.T.T., Ellwood, S.R., Mathur, P.N., Sarker, A., and Ford, R. (2012a). Integration of EST-SSR markers of *Medicago truncatula* into intraspecific linkage map of lentil and identification of QTL conferring resistance to Ascochyta Blight at seedling and pod stages. Molecular Breeding *30*, 429-439.

Gupta, M., Verma, B., Kumar, N., Chahota, R., Rathour, R., Sharma, S., Bhatia, S., and Sharma, T. (2012b). Construction of intersubspecific molecular genetic map of lentil based on ISSR, RAPD and SSR markers. J Genet *91*, 279-287.

Hamwieh, A., Udupa, S.M., Choumane, W., Sarker, A., Dreyer, F., Jung, C., and Baum, M. (2005). A genetic linkage map of *Lens* sp. based on microsatellite and AFLP markers and the localization of Fusarium Vascular Wilt resistance. Theor Appl Genet *110*, 669-677.

Harlan, J.R., de Wet, J.M.J., and Price, E.G. (1973). Comparative evolution of cereals. Evolution *27*, 311-325.

Hawtin, G.C., Muehlbauer, F.J., Slinkard, A.E., and Singh, K.B. (1988). Current status of cool season food legume crop improvement: an assessment of critical needs. In World crops: Cool season food legumes: A global perspective of the problems and prospects for crop improvement in pea, lentil, faba bean and chickpea, R.J. Summerfield, ed. (Springer Netherlands).

Hecht, V., Foucher, F., Ferrándiz, C., Macknight, R., Navarro, C., Morin, J., Vardy, M.E., Ellis, N., Beltrán, J.P., Rameau, C*., et al.* (2005). Conservation of *Arabidopsis* flowering genes in model legumes. Plant Physiology *137*, 1420-1434.

Hecht, V., Knowles, C.L., Vander Schoor, J.K., Liew, L.C., Jones, S.E., Lambert, M.J.M., and Weller, J.L. (2007). Pea *LATE BLOOMER1* is a *GIGANTEA* ortholog with roles in photoperiodic flowering, deetiolation, and transcriptional regulation of circadian clock gene homologs. Plant Physiology *144*, 648-661.

Hecht, V., Laurie, R.E., Vander Schoor, J.K., Ridge, S., Knowles, C.L., Liew, L.C., Sussmilch, F.C., Murfet, I.C., Macknight, R.C., and Weller, J.L. (2011). The Pea *GIGAS* Gene Is a *FLOWERING LOCUS T* Homolog Necessary for Graft-Transmissible Specification of Flowering but Not for Responsiveness to Photoperiod. The Plant Cell Online *23*, 147-161.

Helfer, A., Nusinow, D.A., Chow, B.Y., Gehrke, A.R., Bulyk, M.L., and Kay, S.A. (2011). *LUX ARRHYTHMO* Encodes a Nighttime Repressor of Circadian Gene Expression in the *Arabidopsis* Core Clock. Current Biology *21*, 126-133.

Heo, J.B., and Sung, S. (2011). Vernalization-mediated epigenetic silencing by a long intronic noncoding RNA. Science *331*, 76-79.

Herrero, E., Kolmos, E., Bujdoso, N., Yuan, Y., Wang, M., Berns, M.C., Uhlworm, H., Coupland, G., Saini, R., Jaskolski, M*., et al.* (2012). *EARLY FLOWERING 4* recruitment of *EARLY FLOWERING 3* in the nucleus sustains the *Arabidopsis* circadian clock. The Plant Cell Online.

Ito, S., Song, Y.H., Josephson-Day, A.R., Miller, R.J., Breton, G., Olmstead, R.G., and Imaizumi, T. (2012). *FLOWERING BHLH* transcriptional activators control expression of the photoperiodic flowering regulator *CONSTANS* in *Arabidopsis*. Proceedings of the National Academy of Sciences of the United States of America *109*, 3582-3587.

Jaudal, M., Yeoh, C.C., Zhang, L., Stockum, C., Mysore, K.S., Ratet, P., and Putterill, J. (2013). Retroelement insertions at the *Medicago FTa1* locus in spring mutants eliminate vernalisation but not long-day requirements for early flowering. The Plant Journal *76*, 580-591.

Kahraman, A., Kusmenoglu, I., Aydin, N., Aydogan, A., Erskine, W., and Muehlbauer, F.J. (2004). QTL mapping of winter hardiness genes in lentil. Crop Science *44*, 13-22.

Kahriman, A., Temel, H.Y., Aydogan, A., and Tanyolac, M.B. (2014). Major quantitative trait loci for flowering time in lentil. Turkish Journal of Agriculture and Forestry, 588.

Kamphuis, L.G., Williams, A.H., D'Souza, N.K., Pfaff, T., Ellwood, S.R., Groves, E.J., Singh, K.B., Oliver, R.P., and Lichtenzveig, J. (2007). The *Medicago truncatula* reference accession A17 has an aberrant chromosomal configuration. New Phytologist *174*, 299-303.

Kardailsky, I., Shukla, V.K., Ahn, J.H., Dagenais, N., Christensen, S.K., Nguyen, J.T., Chory, J., Harrison, M.J., and Weigel, D. (1999). Activation Tagging of the Floral Inducer *FT*. Science *286*, 1962-1965.

Kato, K., and Yokoyama, H. (1992). Geographical variation in heading characters among wheat landraces, *Triticum aestivum* L., and its implication for their adaptability. Theor Appl Genet *84*, 259-265.

Kaur, S., Cogan, N.I., Stephens, A., Noy, D., Butsch, M., Forster, J., and Materne, M. (2014). EST-SNP discovery and dense genetic mapping in lentil (*Lens culinaris* Medik.) enable candidate gene selection for boron tolerance. Theor Appl Genet *127*, 703-713.

Kaur, S., Cogan, N.O., Pembleton, L.W., Shinozuka, M., Savin, K.W., Materne, M., and Forster, J.W. (2011). Transcriptome sequencing of lentil based on second-generation technology permits large-scale unigene assembly and SSR marker discovery. BMC Genomics *12*.

Kobayashi, Y., Kaya, H., Goto, K., Iwabuchi, M., and Araki, T. (1999). A Pair of Related Genes with Antagonistic Roles in Mediating Flowering Signals. Science *286*, 1960-1962.

Kolmos, E., Herrero, E., Bujdoso, N., Millar, A.J., Tóth, R., Gyula, P., Nagy, F., and Davis, S.J. (2011). A reduced-function allele reveals that *EARLY FLOWERING 3* repressive action on the circadian clock is modulated by phytochrome signals in *Arabidopsis*. The Plant Cell Online.

Kumar, Y., Mishra, S., Tyagi, M., Singh, S., and Sharma, B. (2005). Linkage between genes for leaf colour, plant pubescence, number of leaflets and plant height in lentil (*Lens culinaris* Medik.). Euphytica *145*, 41-48.

Ladizinsky, G. (1979). The genetics of several morphological traits in the lentil. Journal of Heredity *70*, 135-137.

Ladizinsky, G. (1985). The genetics of hard seed coat in the genus *Lens*. Euphytica *34*, 539-543.

Ladizinsky, G., Braun, D., Goshen, D., and Muehlbauer, F.J. (1984). The biological species of the genus *Lens* L. Botanical gazette *145*, 253-261.

Ladizinsky, G., Cohen, D., and Muehlbauer, F. (1985). Hybridization in the genus *Lens* by means of embryo culture. Theor Appl Genet *70*, 97-101.

Laurie, R.E., Diwadkar, P., Jaudal, M., Zhang, L., Hecht, V., Wen, J., Tadege, M., Mysore, K.S., Putterill, J., Weller, J.L*., et al.* (2011). The *Medicago FLOWERING LOCUS T* homolog, *MtFTa1*, is a key regulator of flowering time. Plant Physiology *156*, 2207-2224.

Le Masson, I., Jauvion, V., Bouteiller, N., Rivard, M., Elmayan, T., and Vaucheret, H. (2012). Mutations in the *Arabidopsis* H3K4me2/3 demethylase *JMJ14* suppress posttranscriptional gene silencing by decreasing transgene transcription. The Plant Cell *24*, 3603-3612.

Liew, L.C., Hecht, V., Laurie, R.E., Knowles, C.L., Vander Schoor, J.K., Macknight, R.C., and Weller, J.L. (2009a). *DIE NEUTRALIS* and *LATE BLOOMER 1* contribute to regulation of the pea circadian clock. The Plant Cell *21*, 3198-3211.

Liew, L.C., Hecht, V., Weeden, N., and Weller, J.L. (2009b). Isolation of *Pseudo Response Regulator* genes and evaluation as candidate genes for photoperiod response loci. Pisum Genetics *41*.

Liew, L.C., Hecht, V.r., Sussmilch, F.C., and Weller, J.L. (2014). The pea photoperiod response gene *STERILE NODES* is an ortholog of *LUX ARRHYTHMO*. Plant Physiology *165*, 648-657.

Liu, A., and Burke, J.M. (2006). Patterns of nucleotide diversity in wild and cultivated sunflower. Genetics *173*, 321-330.

Liu, X., Hao, L., Li, D., Zhu, L., and Hu, S. (2015). Long Non-coding RNAs and Their Biological Roles in Plants. Genomics, Proteomics & Bioinformatics *13*, 137-147.

Liu, X.L., Covington, M.F., Fankhauser, C., Chory, J., and Wagner, D.R. (2001). *ELF3* encodes a circadian clock regulated nuclear protein that functions in an *Arabidopsis PHYB* signal transduction pathway. The Plant Cell Online *13*, 1293-1304.

Lombardi, M., Materne, M., Cogan, N.O., Rodda, M., Daetwyler, H.D., Slater, A.T., Forster, J.W., and Kaur, S. (2014). Assessment of genetic variation within a global collection of lentil (*Lens culinaris* Medik.) cultivars and landraces using SNP markers. BMC genetics *15*, 150.

Lu, F., Cui, X., Zhang, S., Liu, C., and Cao, X. (2010). *JMJ14* is an H3K4 demethylase regulating flowering time in *Arabidopsis*. Cell Res *20*, 387-390.

Lu, S.X., Webb, C.J., Knowles, S.M., Kim, S.H.J., Wang, Z., and Tobin, E.M. (2012). *CCA1* and *ELF3* interact in the control of hypocotyl length and flowering time in *Arabidopsis*. Plant Physiology *158*, 1079-1088.

Manly, K., Cudmore, R., and Meer, J. (2001). Map Manager QTX, cross-platform software for genetic mapping. Mammalian Genome *12*, 930–932.

Matsubara, K., Ogiso-Tanaka, E., Hori, K., Ebana, K., Ando, T., and Yano, M. (2012). Natural variation in *Hd17*, a homolog of *Arabidopsis ELF3* that is involved in rice photoperiodic flowering. Plant and Cell Physiology.

Matsushika, A., Imamura, A., Yamashino, T., and Mizuno, T. (2002). Aberrant expression of the light-inducible and circadian-regulated *APRR9* gene belonging to the circadian-associated *APRR1/TOC1* quintet results in the phenotype of early flowering in *Arabidopsis thaliana*. Plant and Cell Physiology *43*, 833-843.

Mayer, M.S., and Bagga, S.K. (2002). The phylogeny of *Lens* (Leguminosae): new insight from ITS sequence analysis. Plant Systematics and Evolution *232*, 145-154.

McVicar, R., McCall, P., Brenzil, C., Hartley, S., Panchuk, K., Mooleki, P., Vandenberg, A., and Banniza, S. (2010). Lentils in Saskatchewan, S.M.o. Agriculture, ed.

Mockler, T.C., Guo, H., Yang, H., Duong, H., and Lin, C. (1999). Antagonistic actions of *Arabidopsis* cryptochromes and phytochrome B in the regulation of floral induction. Development *126*, 2073-2082.

Muehlbauer, F.J., Kaiser, W.J., Clement, S.L., and Summerfield, R.J. (1995). Production and breeding of Lentil. In Advances in Agronomy, L.S. Donald, ed. (Academic Press), pp. 283-332.

Muehlbauer, F.J., Mihove, M., Vandenberg, A., Tullu, A., and Materne, M. (2009). Improvement in developed countries. In The Lentil: Botany, Production and Uses, W. Erskine, ed. (CABI).

Murfet, I. (1971). Flowering in *Pisum*: reciprocal grafts between known genotypes. Australian Journal of Biological Sciences *24*, 1089-1102.

Nakamichi, N. (2011). Molecular mechanisms underlying the *Arabidopsis* circadian clock. Plant and Cell Physiology.

Nakamichi, N., Kiba, T., Kamioka, M., Suzuki, T., Yamashino, T., Higashiyama, T., Sakakibara, H., and Mizuno, T. (2012). Transcriptional repressor *PRR5* directly regulates clock-output pathways. Proceedings of the National Academy of Sciences *109*, 17123-17128.

Nakamichi, N., Kita, M., Ito, S., Yamashino, T., and Mizuno, T. (2005). *PSEUDO-RESPONSE REGULATORS, PRR9, PRR7* and *PRR5*, together play essential roles close to the circadian clock of *Arabidopsis thaliana*. Plant and Cell Physiology *46*, 686-698.

Nakamichi, N., Kita, M., Niinuma, K., Ito, S., Yamashino, T., Mizoguchi, T., and Mizuno, T. (2007). *Arabidopsis* Clock-Associated Pseudo-Response Regulators *PRR9, PRR7 and PRR5* coordinately and positively regulate flowering time through the canonical *CONSTANS*-dependent photoperiodic pathway. Plant and Cell Physiology *48*, 822-832.

Nefissi, R., Natsui, Y., Miyata, K., Oda, A., Hase, Y., Nakagawa, M., Ghorbel, A., and Mizoguchi, T. (2011). Double loss-of-function mutation in *EARLY FLOWERING 3* and *CRYPTOCHROME 2* genes delays flowering under continuous light but accelerates it under long days and short days: an important role for *Arabidopsis CRY2* to accelerate flowering time in continuous light. Journal of Experimental Botany *62*, 2731-2744.

Nelson, M., Berger, J., and Erskine, W. (2010). Flowering time control in annual legumes: prospects in a changing global climate. CAB Reviews: Perspectives in Agriculture, Veterinary Science, Nutrition and Natural Resources *5*, 1-14.

Nicholas, K., and Nicholas, H. (1997). GeneDoc: a tool for editing and annotating multiple sequence alignments. In GeneDoc.

Nusinow, D.A., Helfer, A., Hamilton, E.E., King, J.J., Imaizumi, T., Schultz, T.F., Farre, E.M., and Kay, S.A. (2011). The *ELF4-ELF3-LUX* complex links the circadian clock to diurnal control of hypocotyl growth. Nature *475*, 398-402.

Panda, S., Poirier, G.G., and Kay, S.A. (2002). tej defines a role for poly(ADP-ribosyl)ation in establishing period length of the *Arabidopsis* circadian oscillator. Developmental Cell *3*, 51-61.

Phan, H.T., Ellwood, S., Hane, J., Ford, R., Materne, M., and Oliver, R. (2007). Extensive macrosynteny between *Medicago truncatula* and *Lens culinaris* ssp. *culinaris*. Theor Appl Genet *114*, 549-558.

Phan, H.T.T., Ellwood, S.R., Ford, R., Thomas, S., and Oliver, R. (2006). Differences in syntenic complexity between *Medicago truncatula* with *Lens culinaris* and *Lupinus albus*. Functional Plant Biology *33*, 775-782.

Purugganan, M.D., and Fuller, D.Q. (2009). The nature of selection during plant domestication. Nature *457*, 843-848.

Rahman, M.M., Sarker, A., Kumar, S., Ali, A., Yadav, N., and Rahman, M.L. (2009). Breeding for short season environments. The lentil: Botany, production and uses, 121.

Rana, M.K., Kumari, R., Singh, S., and Bhat, K.V. (2007). Genetic analysis of Indian Lentil (*Lens culinaris* Medikus) cultivars and landraces using RAPD and STMS markers. J Plant Biochem Biotechnol *16*, 53-57.

Riva, A. (1975). Precoz, a new lentil cultivar for Argentina. Lentil Experimental News Service *2*, 9-10.

Roberts, E.H., Summerfield, R.J., Ellis, R.H., and Stewart, K.A. (1988). Photothermal time for flowering in lentils (*Lens culinaris*) and the analysis of potential vernalization responses. Annals of Botany *61*, 29-39.

Roberts, E.H., Summerfield, R.J., Muehlbauer, F.J., and Short, R.W. (1986). Flowering in lentil (*Lens culinaris* Medic.): the duration of the photoperiodic inductive phase as a function of accumulated daylength above the critical photoperiod. Annals of Botany *58*, 235-248.

Roux, F., Touzet, P., Cuguen, J., and Le Corre, V. (2006). How to be early flowering: an evolutionary perspective. Trends in Plant Science *11*, 375-381.

Rubeena, Ford, R., and Taylor, P.W.J. (2003). Construction of an intraspecific linkage map of lentil (*Lens culinaris* ssp. *culinaris*). Theor Appl Genet *107*, 910-916.

Saha, G.C., Sarker, A., Chen, W., Vandemark, G.J., and Muehlbauer, F.J. (2013). Inheritance and linkage map positions of genes conferring agromorphological traits in *Lens culinaris* Medik. International Journal of Agronomy *2013*, 9.

Salomé, P.A., Bomblies, K., Laitinen, R.A.E., Yant, L., Mott, R., and Weigel, D. (2011). Genetic architecture of flowering-time variation in *Arabidopsis thaliana*. Genetics *188*, 421-433.

Sarker, A., Erskine, W., Sharma, B., and Tyagi, M.C. (1999). Inheritance and linkage relationship of days to flower and morphological loci in lentil (*Lens culinaris* Medikus subsp. *culinaris*). Journal of Heredity *90*, 270-275.

Sato, E., Nakamichi, N., Yamashino, T., and Mizuno, T. (2002). Aberrant expression of the *Arabidopsis* circadian-regulated *APRR5* Gene Belonging to the *APRR1/TOC1* quintet results in early flowering and hypersensitiveness to light in early photomorphogenesis. Plant and Cell Physiology *43*, 1374-1385.

Saxena, M.C. (2009). Plant morphology, anatomy, and growth habit. The lentil: botany, production and uses CABI, Oxfordshire, UK, 34-46.

Serrano, H. (2011). A molecular basis of *ELF3* action in the *Arabidopsis* circadian clock (Universität zu Köln).

Sharma, S., Dawson, I., and Waugh, R. (1995). Relationships among cultivated and wild lentils revealed by RAPD analysis. Theor Appl Genet *91*, 647-654.

Sharpe, A.G., Ramsay, L., Sanderson, L.A., Fedoruk, M.J., Clarke, W.E., Li, R., Kagale, S., Vijayan, P., Vandenberg, A., and Bett, K.E. (2013). Ancient orphan crop joins modern era: gene-based SNP discovery and mapping in lentil. BMC Genomics *14*.

Shrestha, R., Gomez-Ariza, J., Brambilla, V., and Fornara, F. (2014). Molecular control of seasonal flowering in rice, *Arabidopsis* and temperate cereals. Annals of Botany.

Slinkard, A.E. (1988). Indianhead lentil as an annual legume green manure crop for Western Canada. Canadian Journal of Plant Science *68*, 829-841.

Sonnante, G., Hammer, K., and Pignone, D. (2009). From the cradle of agriculture a handful of lentils: history of domestication. Rendiconti Lincei *20*, 21-37.

Soy, J., Leivar, P., and Monte, E. (2014). *PIF1* promotes phytochrome-regulated growth under photoperiodic conditions in *Arabidopsis* together with *PIF3, PIF4*, and *PIF5*. Journal of Experimental Botany *65*, 2925-2936.

Summerfield, R.J., Roberts, E.H., Erskine, W., and Ellis, R.H. (1985). Effects of temperature and photoperiod on flowering in lentils (*Lens culinaris* Medic.). Annals of Botany *56*, 659-671.

Swiezewski, S., Liu, F., Magusin, A., and Dean, C. (2009). Cold-induced silencing by long antisense transcripts of an *Arabidopsis* polycomb target. Nature *462*, 799-802.

Tadmor, Y., Zamir, D., and Ladizinsky, G. (1987). Genetic mapping of an ancient translocation in the genus *Lens*. Theor Appl Genet *73*, 883-892.

Tahir, M., and Muehlbauer, F.J. (1994). Gene mapping in lentil with recombinant inbred lines. Journal of Heredity *85*, 306-310.

Takahashi, Y., and Shimamoto, K. (2011). *Heading date 1* (*Hd*1), an ortholog of *Arabidopsis CONSTANS*, is a possible target of human selection during domestication to diversify flowering times of cultivated rice. Genes & Genetic Systems *86*, 175-182.

Takata, N., Saito, S., Saito, C.T., and Uemura, M. (2010). Phylogenetic footprint of the plant clock system in angiosperms: evolutionary processes of *Pseudo-Response Regulators*. BMC Evolutionary Biology *10*, 126-126.

Tang, H., Krishnakumar, V., Bidwell, S., Rosen, B., Chan, A., Zhou, S., Gentzbittel, L., Childs, K.L., Yandell, M., and Gundlach, H. (2014). An improved genome release (version Mt4.0) for the model legume *Medicago truncatula*. BMC Genomics *15*, 312.

Temel, H.Y., Gol, D., Kahriman, A., and Tanyolac, M.B. (2015). Single nucleotide polymorphism discovery through Illumina-based transcriptome sequencing and mapping in lentil. Turkish Journal of Agriculture and Forestry, 470.

Thompson, J., Gibson, T., Plewniak, F., Jeanmougin, F., and Higgins, D. (1997). The CLUSTALX Windows interface: flexible strategies for multiple sequence alignment aided by quality analysis tools. Nucleic Acids Research *25*, 4876-4882.

Tullu, A., Tar'an, B., Warkentin, T., and Vandenberg, A. (2008). Construction of an intraspecific linkage map and QTL analysis for earliness and plant height in lentil. Crop Sci *48*, 2254-2264.

Turner, A., Beales, J., Faure, S., Dunford, R.P., and Laurie, D.A. (2005). The *Pseudo-Response Regulator Ppd-H1* provides adaptation to photoperiod in barley. Science *310*, 1031-1034.

Van Ooijen, J.W. (2006). JoinMap 4: Software for the calculation of genetic linkage maps in experimental populations (Wageningen, Kyazma, B.V.).

Van Ooijen, J.W. (2009). MapQTL 6: Software for the mapping of quantitative trait loci in experimental populations of diploid species (Wageningen, Kyazma, B.V.).

van Zeist, W. (1970). The Oriental Institute excavations at Mureybi, Syria: preliminary report on the 1965 Campaign Part III: The Paleobotany. Journal of Near Eastern Studies *29*, 167-176.

Vandenberg, A., and Slinkard, A.E. (1989). Inheritance of four new qualitative genes in lentil. Journal of Heredity *80*, 320-322.

Verma, P., Shah, N., and Bhatia, S. (2013). Development of an expressed gene catalogue and molecular markers from the de novo assembly of short sequence reads of the lentil (*Lens culinaris* Medik.) transcriptome. Plant Biotechnology Journal *11*, 894-905.

Vijayan, P., Vandenberg, A., and Bett, K. (2009). A mixed genotype lentil EST library representing the normalized transcriptome of different seed development stages (Saskatoon, University of Saskatchewan).

Wang, S., Basten, C., and Zeng, Z. (2012). Windows QTL Cartographer 2.5 (Raleigh, NC, Department of Statistics, North Carolina State University).

Weeden, N.F., Muehlbauer, F.J., and Ladizinsky, G. (1992). Extensive conservation of linkage relationships between Pea and Lentil genetic maps. Journal of Heredity *83*, 123-129.

Weller, J.L., Batge, S.L., Smith, J.J., Kerckhoffs, L.H.J., Sineshchekov, V.A., Murfet, I.C., and Reid, J.B. (2004). A dominant mutation in the pea *PHYA* gene confers enhanced responses to light and impairs the light-dependent degradation of phytochrome A. Plant Physiology *135*, 2186-2195.

Weller, J.L., Hecht, V., Liew, L.C., Sussmilch, F.C., Wenden, B., Knowles, C.L., and Vander Schoor, J.K. (2009). Update on the genetic control of flowering in garden pea. Journal of Experimental Botany *60*, 2493-2499.

Weller, J.L., Liew, L.C., Hecht, V.F.G., Rajandran, V., Laurie, R.E., Ridge, S., Wenden, B., Vander Schoor, J.K., Jaminon, O., Blassiau, C.*, et al.* (2012). A conserved molecular basis for photoperiod adaptation in two temperate legumes. Proceedings of the National Academy of Sciences *109*, 21158-21163.

Weller, J.L., and Murfet, I.C. (unpublished). Characterisation of lentil *Sn* (Hobart, University of Tasmania).

Weller, J.L., and Ortega Martinez, R. (2015). Genetic control of flowering time in legumes. Frontiers in Plant Science *6*.

Wong, A.C.S., Hecht, V.F.G., Picard, K., Diwadkar, P., Laurie, R.E., Wen, J., Mysore, K., Macknight, R.C., and Weller, J.L. (2014). Isolation and functional analysis of *CONSTANS-LIKE* genes suggests that a central role for *CONSTANS* in flowering time control is not evolutionarily conserved in *Medicago truncatula*. Frontiers in Plant Science *5*, 486.

Wu, W., Zheng, X.-M., Lu, G., Zhong, Z., Gao, H., Chen, L., Wu, C., Wang, H.-J., Wang, Q., Zhou, K.*, et al.* (2013). Association of functional nucleotide polymorphisms at *DTH2* with the northward expansion of rice cultivation in Asia. Proceedings of the National Academy of Sciences *110*, 2775-2780.

Yamamoto, Y., Sato, E., Shimizu, T., Nakamich, N., Sato, S., Kato, T., Tabata, S., Nagatani, A., Yamashino, T., and Mizuno, T. (2003). Comparative Genetic Studies on the *APRR5* and *APRR7* Genes Belonging to the *APRR1/TOC1* Quintet Implicated in Circadian Rhythm, Control of Flowering Time, and Early Photomorphogenesis. Plant and Cell Physiology *44*, 1119-1130.

Yano, M., Katayose, Y., Ashikari, M., Yamanouchi, U., Monna, L., Fuse, T., Baba, T., Yamamoto, K., Umehara, Y., Nagamura, Y.*, et al.* (2000). *Hd1*, a major photoperiod sensitivity quantitative trait locus in rice, is closely related to the *Arabidopsis* flowering time gene *CONSTANS*. The Plant Cell *12*, 2473-2483.

Yeoh, C.C., Balcerowicz, M., Zhang, L., Jaudal, M., Brocard, L., Ratet, P., and Putterill, J. (2013). Fine Mapping Links the *FTa1* Flowering Time Regulator to the Dominant *Spring1* Locus in *Medicago*. PLoS ONE *8*, e53467.

Yu, J.-W., Rubio, V., Lee, N.-Y., Bai, S., Lee, S.-Y., Kim, S.-S., Liu, L., Zhang, Y., Irigoyen, M.L., Sullivan, J.A*., et al.* (2008). *COP1* and *ELF3* control circadian function and photoperiodic flowering by regulating *GI* stability. Molecular cell *32*, 617-630.

Zagotta, M.T., Hicks, K.A., Jacobs, C.I., Young, J.C., Hangarter, R.P., and Meeks-Wagner, D.R. (1996). The *Arabidopsis ELF3* gene regulates vegetative photomorphogenesis and the photoperiodic induction of flowering. The Plant Journal *10*, 691-702.

Zamir, D., and Ladizinsky, G. (1984). Genetics of allozyme variants and linkage groups in lentil. Euphytica *33*, 329-336.

Zentner, R.P., Campbell, C.A., Biederbeck, V.O., and Selles, F. (1996). Indianhead black lentil as green manure for wheat rotations in the Brown soil zone. Canadian Journal of Plant Science *76*, 417-422.

Zohary, D. (1972). The wild progenitor and the place of origin of the cultivated lentil: *Lens culinaris*. Econ Bot *26*, 326-332.

Zohary, D. (2004). Unconscious selection and the evolution of domesticated plants. Econ Bot *58*, 5-10.

Zohary, D., and Hopf, M. (1973). Domestication of pulses in the old world: legumes were companions of wheat and barley when agriculture began in the Near East. Science *182*, 887-894.

Appendix

Appendix 1 *University of Saskatchewan Lentil Association Mapping (LAM) panel*

S/N	Accession	Taxon	Country	Latitude	Longitude	Altitude (m)
1	ILL28	L. culinaris subsp. culinaris	Syria	35.6	36.7	475
2	ILL55	L. culinaris subsp. culinaris	Iraq	36.7	43.3	568
3	ILL132	L. culinaris subsp. culinaris	Turkey	-	-	-
4	ILL141	L. culinaris subsp. culinaris	Turkey	39.1	39.6	1206
5	ILL229	L. culinaris subsp. culinaris	Pakistan	34	71.7	297
6	ILL293	L. culinaris subsp. culinaris	Greece	38.7	20.7	547
7	ILL313	L. culinaris subsp. culinaris	Israel	32.8	35	124
8	ILL572	L. culinaris subsp. culinaris	Turkey	38.8	39.5	1070
9	ILL1048	L. culinaris subsp. culinaris	Iran	29.1	54.1	2077
10	ILL1139	L. culinaris subsp. culinaris	Lebanon	-	-	-
11	ILL1220	L. culinaris subsp. culinaris	Iran	29.6	52.5	1598
12	ILL1462	L. culinaris subsp. culinaris	Iran	28.7	57.8	877
13	ILL1553	L. culinaris subsp. culinaris	Iran	32.7	51.7	1558
14	ILL1671	L. culinaris subsp. culinaris	Azerbaijan	39.4	45.2	892
15	ILL1704	L. culinaris subsp. culinaris	Ethiopia	9	38.4	2453
16	ILL1744	L. culinaris subsp. culinaris	Ethiopia	8.3	37.7	1789
17	ILL2433	L. culinaris subsp. culinaris	Ethiopia	5.3	39.6	1462
18	ILL2501	L. culinaris subsp. culinaris	India	-	-	-
19	ILL2607	L. culinaris subsp. culinaris	India	-	-	-
20	ILL3025	L. culinaris subsp. culinaris	India	-	-	-
21	ILL3167	L. culinaris subsp. culinaris	India	-	-	-
22	ILL3347	L. culinaris subsp. culinaris	India	-	-	-
23	ILL3487	L. culinaris subsp. culinaris	Nepal	26.9	86.1	181
24	ILL3502	L. culinaris subsp. culinaris	India	29.7	72.4	200
25	ILL3596	L. culinaris subsp. culinaris	India	34.7	72.2	790
26	ILL3597	L. culinaris subsp. culinaris	India	-	-	-
27	ILL3714	L. culinaris subsp. culinaris	India	-	-	-
28	ILL3805	L. culinaris subsp. culinaris	India	-	-	-
29	ILL3925	L. culinaris subsp. culinaris	India	-	-	-
30	ILL4080	L. culinaris subsp. culinaris	India	-	-	-
31	ILL4164	L. culinaris subsp. culinaris	India	37.4	36.2	480
32	ILL4349	L. culinaris subsp. culinaris	Canada	-	-	-
33	ILL4359	L. culinaris subsp. culinaris	India	-	-	-
34	ILL4387	L. culinaris subsp. culinaris	Egypt	-	-	-
35	ILL4400	L. culinaris subsp. culinaris	Syria	-	-	-

Continued on next page

Continued from previous page

S/N	Accession	Taxon	Country	Latitude	Longitude	Altitude (m)
36	ILL4542	L. culinaris subsp. culinaris	Syria	37.1	41.7	531
37	ILL4605	L. culinaris subsp. culinaris	Argentina			
38	ILL4609	L. culinaris subsp. culinaris	Pakistan	-	-	-
39	ILL4665	L. culinaris subsp. culinaris	Hungary	-	-	-
40	ILL4671	L. culinaris subsp. culinaris	USA	-	-	-
41	ILL4740	L. culinaris subsp. culinaris	France	47.4	1.6	106
42	ILL4768	L. culinaris subsp. culinaris	Yemen	15.2	44	2800
43	ILL4774	L. culinaris subsp. culinaris	Romania	-	-	-
44	ILL4778	L. culinaris subsp. culinaris	Uruguay	-	-	-
45	ILL4782	L. culinaris subsp. culinaris	Norway	-	-	-
46	ILL4783	L. culinaris subsp. culinaris	Czech Republic	-	-	-
47	ILL4785	L. culinaris subsp. culinaris	Slovakia	-	-	-
48	ILL4804	L. culinaris subsp. culinaris	Libya	-	-	-
49	ILL4831	L. culinaris subsp. culinaris	Germany	-	-	-
50	ILL4865	L. culinaris subsp. culinaris	Greece	37.9	22.3	1201
51	ILL4875	L. culinaris subsp. culinaris	Unknown	-	-	-
52	ILL4915	L. culinaris subsp. culinaris	Croatia	-	-	-
53	ILL4956	L. culinaris subsp. culinaris	Portugal	-	-	-
54	ILL5058	L. culinaris subsp. culinaris	Spain	40.9	-1.9	1176
55	ILL5151	L. culinaris subsp. culinaris	India	22.6	88.4	6
56	ILL5209	L. culinaris subsp. culinaris	Jordan	32.6	35.9	515
57	ILL5372	L. culinaris subsp. culinaris	Jordan	30.6	35.6	1102
58	ILL5399	L. culinaris subsp. culinaris	Bulgaria	-	-	-
59	ILL5418	L. culinaris subsp. culinaris	Italy	41.3	15.2	1450
60	ILL5425	L. culinaris subsp. culinaris	Mexico	-	-	-
61	ILL5490	L. culinaris subsp. culinaris	Chile	-	-	-
62	ILL5511	L. culinaris subsp. culinaris	Syria	37.1	41.3	515
63	ILL5576	L. culinaris subsp. culinaris	Serbia	44.8	20.5	129
64	ILL5584	L. culinaris subsp. culinaris	Jordan	31.3	35.7	893
65	ILL5588	L. culinaris subsp. culinaris	Jordan	32.1	35.7	700
66	ILL5883	L. culinaris subsp. culinaris	Jordan	-	-	-
67	ILL5945	L. culinaris subsp. culinaris	Ethiopia	-	-	-
68	ILL6166	L. culinaris subsp. culinaris	Unknown	-	-	-
69	ILL6182	L. culinaris subsp. culinaris	Tunisia	-	-	-
70	ILL6264	L. culinaris subsp. culinaris	Unknown	-	-	-

Continued on next page

Continued from previous page

S/N	Accession	Taxon	Country	Latitude	Longitude	Altitude (m)
71	ILL6378	L. culinaris subsp. culinaris	Unknown	-	-	-
72	ILL6505	L. culinaris subsp. culinaris	Unknown	-	-	-
73	ILL6540	L. culinaris subsp. culinaris	Unknown	-	-	-
74	ILL6689	L. culinaris subsp. culinaris	Unknown	-	-	-
75	ILL6853	L. culinaris subsp. culinaris	Syria	-	-	-
76	ILL6920	L. culinaris subsp. culinaris	Unknown	-	-	-
77	ILL6967	L. culinaris subsp. culinaris	Brazil	-	-	-
78	ILL7051	L. culinaris subsp. culinaris	Algeria	-	-	-
79	ILL7089	L. culinaris subsp. culinaris	Russia	-	-	-
80	ILL7499	L. culinaris subsp. culinaris	Unknown	-	-	-
81	ILL7567	L. culinaris subsp. culinaris	Jordan	32.2	36.2	700
82	ILL7585	L. culinaris subsp. culinaris	Turkey	-	-	-
83	ILL7621	L. culinaris subsp. culinaris	Iran	-	-	-
84	ILL7727	L. culinaris subsp. culinaris	Unknown	-	-	-
85	ILL7745	L. culinaris subsp. culinaris	Unknown	-	-	-
86	ILL7747	L. culinaris subsp. culinaris	Syria	36.8	36.7	500
87	ILL7773	L. culinaris subsp. culinaris	Unknown	-	-	-
88	ILL7791	L. culinaris subsp. culinaris	Unknown	-	-	-
89	ILL9901	L. culinaris subsp. culinaris	Unknown	-	-	-
90	964a-46	L. culinaris subsp. culinaris	Unknown	-	-	-
91	Eston	L. culinaris subsp. culinaris	Canada	-	-	-
92	Indianhead	L. culinaris subsp. culinaris	Canada	-	-	-
93	PI320937	L. culinaris subsp. culinaris	Germany	-	-	-
94	CDC Robin	L. culinaris subsp. culinaris	Canada	-	-	-

End of Appendix 1

Appendix 2 *Summary of molecular markers and qPCR primers*

	Name	Medicago locus (Mt4.0)	Population	Marker type	Forward Primer	Reverse Primer	Comments	Source
Chapter 3	*MYB1*	Medtr7g146190	ILL 6005 x ILL 5588	CAPS	GGTAGTGGCTGAGATTGG	AACCACTTCTTCACCAGAGG		This study
	ELF3	Medtr3g103970	ILL 6005 x ILL 5588	CAPS	GCAGTAATAGGCCAAAAACATTTCGG	CGATCCGGCAATTAGTTGTT	*elf3-1*	This study
			ILL 6005 x ILL 5588, ILL 223 x ILL 5588	HRM	GGACAGGCAGTCCAAAACAG	GTTGACAAATTTTCTACCCTGGAG	*elf3-2*	This study
			ILL 6005 x ILL 5588	KASP	GAAGGTGACCAAGTTCATGCTGCACTGATCAATGTCAAAGATAATTTACC	AGAGTGTTTGCAGTACAAGTGTTTGAGTT	*elf3-1*	This study
					GAAGGTCGGAGTCAACGGATTAGCACTGATCAATGTCAAAGATAATTTACT	GTACAAGTGTTTGAGTTGCATAGACTGAT		
			ILL 6005 x ILL 5588	KASP	GAAGGTGACCAAGTTCATGCTACCCTGGAGATCTTGGAAATGTTGTC	GGGACAGGCAGTCCAAAACAGTTA	*elf3-2*	This study
					GAAGGTCGGAGTCAACGGATTCTACCCTGGAGATCTTGGAAATGTTATT	CAGGCAGTCCAAAACAGTTAGGAAATTTA		
	LUX	Medtr4g064730	ILL 6005 x ILL 5588	CAPS	CAAGAGATTCGTCGACGTTGTGG	CTCAAGAATCAGTTTGTGCA		This study
	TOC1	unknown	ILL 6005 x ILL 5588	CAPS	TTCTCTGATGACACAGACGAC	GCAGCAGTGGCAATACTAGC		This study
	PRR37	Medtr4g079920	ILL 6005 x ILL 5588	CAPS	TGGCAACATGTTTGGAGAAG	ACACTCAAGCCTCTGCTTCC		This study
	PRR59	Medtr3g092780	ILL 6005 x ILL 5588	CAPS	GGTATCTGGCTATGCACTTCTCTCG	CAAACATGCTGCCACAGATT		This study
Chapter 5	*EF1a*	Medtr1g101880	-	qPCR	GATGCACCTGGACATCGTGAC	CTTAGGGGTGGTAGCATCCATCT		Johnston
	FTa1-FTa2	Medtr7g084970 - Medtr7g085020 intergenic region	ILL 2601 x ILL 5588	Allele-specific PCR	TGGGCTTGATACTTTGTACTCC TCTACACACTTTGCTGGTTTTG	CCATCACAATTCAAAGCAATG	Figure 5-5C	This study
	MYB1	Medtr7g146190	ILL 2601 x ILL 5588	HRM	TGGACCAAATCTTAATTGGATCTC	TGCACAGCAGTTTTTGTTCC		This study
	PRR59c	Medtr7g118260	ILL 2601 x ILL 5588	HRM	TTTGAAGTGAAAACGTTCAAGT	TGCGTGCTTAAAATGAATCAA		This study
	PIF3c	Medtr7g110810	ILL 2601 x ILL 5588	HRM	CTTCATGCCGTTTTCTATGT	TGGCATACAAAGTCCTCTACCC		This study
	COLg	Medtr7g108150	ILL 2601 x ILL 5588	HRM	CCGTATCAGAGTTGGCACTG	TCAGTTCCTATGAAACACAGACA		This study
	FTa1	Medtr7g084970	-	HRM	GATCAGTTTCGGTACGTACATTTT	AAATAGCTGAATTTTCACTTTCAT		This study
	FTa1	Medtr7g084970	-	qPCR	CCGATATTCCAGCAACTACTGA	AACACGAACACGAAACGATG		This study
	FTa2	Medtr7g085020	-	qPCR	CGGAAATAGGAATGTTTCCAATGG	AACTTGGGCTAGGTGCATCA		This study
	FTc	Medtr7g085040	-	qPCR	GATATTCCAGCCACAACAAGC	CAAACAACTGGGCAGAGACA		This study

Appendix 3 *Primer information*

	Name	Medicago locus (Mt4.0)	Forward Primer	Reverse Primer	Comments
Chapter 3	*MYB1*	Medtr7g146190	GGCTTTACTATTTGCTGTGC	GAATTTTTGCGCATGGCTCCTG	partial genomic sequence isolation
	ELF3	Medtr3g103970	TGGACATGGACAAAGTGACG	CGATCCGGCAATTAGTTGTT	exon 3 splicing defect (Figure 3-5C)
			GTTTAGAGTTTAGGATAGAAAAGGGGTAGG	GCAATTTCTTTTCTGGCTTTCC	full length genomic sequence isolation
			TGGACATGGACAAAGTGACG	CGATCCGGCAATTAGTTGTT	
			TGTTTGCAGTCCAAGTGTTTG	CCGTTACATGATGGCACACC	
			GTTTAGAGTTTAGGATAGAAAAGGGGTAGG	CGATCCGGCAATTAGTTGTT	full length coding sequence isolation
			TGGACATGGACAAAGTGACG	CCGTTACATGATGGCACACC	
	LUX	Medtr4g064730	CAAGAGATTCGTCGACGTTGTGG	CTCAAGAATCAGTTTGTGCA	partial isolation
	TOC1	unknown	TTCTCTGATGACACAGACGAC	GCAGCAGTGGCAATACTAGC	partial isolation
	PRR37	Medtr4g079920	TGGCAACATGTTTGGAGAAG	ACACTCAAGCCTCTGCTTCC	partial isolation
	PRR59b	Medtr3g092780	GGTATCTGGCTATGCACTTCTCTCG	CAAACATGCTGCCACAGATT	partial isolation
Chapter 5	*FTa1*	Medtr7g084970	CCACATATGGCAGGTAGTAGC	TGCATATAACTAGTGCTTGCTTG	partial isolation
	FTa2	Medtr7g085020	GGAAATGACCCCGTGATCTA	TTGCTGGAATATCAGTCACCATCC	partial isolation
	FTa1-FTa2 intergenic	-	ACCTCAGGGATCCATCGTTT	TTACACGCCCAACAACAAGA	full length genomic sequence isolation (ILL 2601)
	FTc	Medtr7g085040	ATGCCTAGGAATATGGTCGATCC	CTTGCGCTTGTTGTGGCTGG	partial isolation
	PRR59c	Medtr7g118260	CACCAATCTGCGATTTAGGG	AGGCTGCAACAACATACTGC	full length coding sequence isolation, full length coding sequence isolation
			CCTGGCAGCTCATCTAAACC	TCAAACAGATTTCACCAGACC	
	PIF3c	Medtr7g110810	GAAAGAGCTCATACCAAATTGC	TGGCATACAAAGTCCTCTACCC	partial isolation
	COLg	Medtr7g108150	GAAGAAAGTTTTGGGGATGG	TGGAATTTCTTCGGCATGAG	partial isolation
	MYB1	Medtr7g146190	GGCTTTACTATTTGCTGTGC	GAATTTTTGCGCATGGCTCCTG	partial isolation

Appendix 4 *Sequence information*

Species	Gene Symbol	Locus ID/Scaffold	Source
Arabidopsis thaliana	*APRR5*	AT5G24470.1	https://www.arabidopsis.org/
Arabidopsis thaliana	*APRR9*	AT2G46790.1	
Arabidopsis thaliana	*ELF3*	AT2G25930.1	
Brachypodium distachyon	*BdPRR59*	Bradi4g24967.1	http://phytozome.jgi.doe.gov/
Brachypodium distachyon	*BdPRR95*	Bradi4g36077.1	
Cicer arietinum	*CaPRR59a*	XP_004502760	http://cicar.comparative-legumes.org/
Cicer arietinum	*CaPRR59b*	*sequence not available**	
Cicer arietinum	*CaPRR59c*	XP_004494872	
Cicer arietinum	*CaELF3*	XP_004501482.1	
Glycine max	*GmPRR59a1*	Glyma04g40640.1	http://phytozome.jgi.doe.gov/
Glycine max	*GmPRR59a2*	Glyma06g14150.2	
Glycine max	*GmPRR59b1*	Glyma07g05530.2	
Glycine max	*GmPRR59b2*	Glyma16g02050.2	
Glycine max	*GmPRR59c1*	Glyma03g42221.1	
Glycine max	*GmELF3*	Glyma19g44970.1	
Glycine max	*GmPRR59c3*	Glyma04g05280	
Lens culinaris spp. culinaris	*LcPRR59a*	Lc0.7_scaffold47127-1	http://knowpulse.usask.ca/portal/
Lens culinaris spp. culinaris	*LcPRR59b*	Lc0.7_scaffold55821-1	
Lens culinaris spp. culinaris	*LcPRR59c*	Lc0.7_scaffold37819-1	
Lens culinaris spp. culinaris	*Lcprr59c*	-	This study
Lens culinaris spp. culinaris	*LcELF3*	JX946295.1	This study
Lens culinaris spp. culinaris	*Lcelf3-1*	-	
Lotus japonicus	*LjPRR59a*	Lj1g3v1076690.1	http://www.kazusa.or.jp/
Lotus japonicus	*LjPRR59b*	chr1.CM0105.1590.r2.m	
Lotus japonicus	*LjPRR59c*	chr3.CM0208.230.r2.m	
Medicago truncatula	*MtPRR59a*	Medtr3g092780.1	http://jcvi.org/medicago/
Medicago truncatula	*MtPRR59b*	Medtr8g024260.1	
Medicago truncatula	*MtPRR59c*	Medtr7g118260.1	
Medicago truncatula	*MtELF3*	Medtr3g103970.1	
Oryza sativa	*OsPRR59*	Os11g05930.1	http://phytozome.jgi.doe.gov/
Oryza sativa	*OsPRR95*	Os09g36220.1	
Phaseolus vulgaris	*PvPRR59a*	Phvul009g045000.1	
Phaseolus vulgaris	*PvPRR59b*	Phvul010g119700.1	
Phaseolus vulgaris	*PvPRR59c*	Phvul009g258300.1	
Pisum sativum	*PsELF3/HR*	JN983406.1	Weller et al. (2012)
Populus trichocarpa	*PtPRR59a1*	Potri.012G005900.1	http://phytozome.jgi.doe.gov/
Populus trichocarpa	*PtPRR59a2*	Potri.015G002300.1	
Populus trichocarpa	*PtPRR59b1*	Potri.002G179800.1	
Populus trichocarpa	*PtPRR59b2*	Potri.014G106000.1	

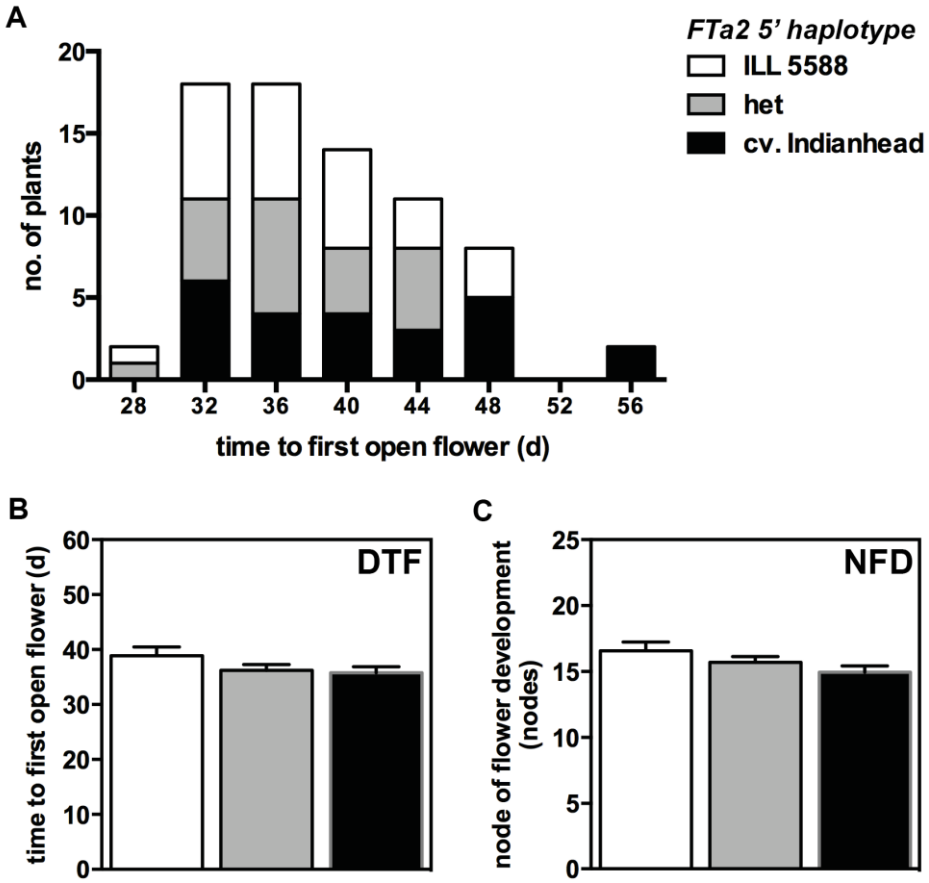
Asterisk (*) *CaPRR59b* is predicted to be positioned between two scaffolds in the International Chickpea Genome Sequencing Consortium (ICGSC) assembly v1.0.

.

Appendix 5 *University of Tasmania Lentil Collection*

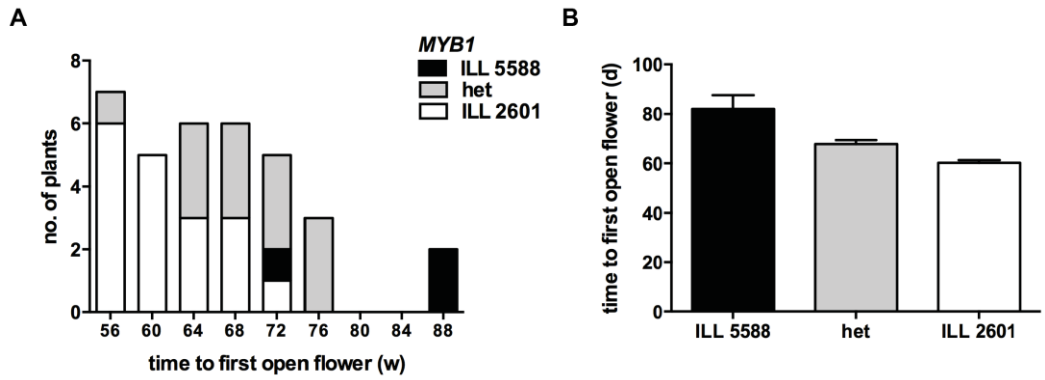
S/N	Accession	Taxon	Country	Latitude	Longitude	Altitude (m)
1	ILL 1823	*L. culinaris subsp. culinaris*	Afghanistan	36.9	70.9	1525
2	ILL 2153	*L. culinaris subsp. culinaris*	Iraq	36.8	43.0	460
3	ILL 2214	*L. culinaris subsp. culinaris*	Lebanon	33.8	36.0	910
4	ILL 223	*L. culinaris subsp. culinaris*	Iran	38.7	46.3	1360
5	ILL 2276	*L. culinaris subsp. culinaris*	Lebanon	33.5	35.4	30
6	ILL 2601	*L. culinaris subsp. culinaris*	India	23.0	-	-
7	ILL 4349	*L. culinaris subsp. culinaris*	Canada	-	-	-
8	ILL 4370	*L. culinaris subsp. culinaris*	Iraq	33.3	44.4	40
9	ILL 4605	*L. culinaris subsp. culinaris*	Argentina	-27.0	-	-
10	ILL 5065	*L. culinaris subsp. culinaris*	Jordan	32.6	35.7	300
11	ILL 5588	*L. culinaris subsp. culinaris*	Jordan	32.1	35.7	700
12	ILL 5895	*L. culinaris subsp. culinaris*	Ethiopia	7.1	39.1	2440
13	ILL 5976	*L. culinaris subsp. culinaris*	Cyprus	34.8	33.0	260
14	ILL 6005	*L. culinaris subsp. culinaris*	Argentina	-27.0	-	-
15	Indianhead	*L. culinaris subsp. culinaris*	Canada	-	-	-
16	PI 297774	*L. culinaris subsp. culinaris*	Greece	40.9	23.7	140
17	PI 297779	*L. culinaris subsp. culinaris*	Greece	39.7	20.8	580
18	PI 297789	*L. culinaris subsp. culinaris*	Greece	38.3	20.5	700
19	PI 298122	*L. culinaris subsp. culinaris*	France	48.3	-1.2	60
20	PI 300248	*L. culinaris subsp. culinaris*	Syria	36.0	37.2	400
21	PI 300250	*L. culinaris subsp. culinaris*	Syria	36.4	37.2	460
22	PI 339281	*L. culinaris subsp. culinaris*	Turkey	38.2	29.0	780
23	PI 339293	*L. culinaris subsp. culinaris*	Turkey	38.3	33.6	920
24	PI 345552	*L. culinaris subsp. culinaris*	Russia	42.0	47.0	3200
25	PI 345631	*L. culinaris subsp. culinaris*	Russia	53.2	45.0	220
26	PI 357227	*L. culinaris subsp. culinaris*	Macedonia	42.0	22.2	450
27	PI 368649	*L. culinaris subsp. culinaris*	Serbia	45.4	19.2	620
28	PI 374116	*L. culinaris subsp. culinaris*	Morocco	31.8	-8.3	480
29	PI 420924	*L. culinaris subsp. culinaris*	Jordan	32.1	35.8	860
30	PI 426775	*L. culinaris subsp. culinaris*	Pakistan	27.6	68.1	42
31	PI 426788	*L. culinaris subsp. culinaris*	Pakistan	31.9	73.1	180
32	PI 426797	*L. culinaris subsp. culinaris*	Pakistan	32.3	74.4	234
33	PI 472291	*L. culinaris subsp. culinaris*	India	26.7	85.2	50
34	PI 472311	*L. culinaris subsp. culinaris*	India	22.4	88.1	10
35	PI 472317	*L. culinaris subsp. culinaris*	India	14.0	77.0	600
36	PI 472328	*L. culinaris subsp. culinaris*	India	27.6	81.6	126
37	PI 472343	*L. culinaris subsp. culinaris*	India	26.3	78.2	180
38	PI 472359	*L. culinaris subsp. culinaris*	India	26.0	93.0	60
39	PI 472578	*L. culinaris subsp. culinaris*	Iran	35.8	51.0	1300
40	PI 472625	*L. culinaris subsp. culinaris*	Iran	35.2	59.4	1340
41	PI 509333	*L. culinaris subsp. culinaris*	Turkey	37.2	38.8	650
42	PI 509409	*L. culinaris subsp. culinaris*	Turkey	39.7	35.8	1300
43	PI 513328	*L. culinaris subsp. culinaris*	Pakistan	25.3	68.8	20
44	PI 533693	*L. culinaris subsp. culinaris*	Spain	42.3	-5.4	860
45	PI 606600	*L. culinaris subsp. culinaris*	Nepal	29.6	81.3	1368
46	PI 606610	*L. culinaris subsp. culinaris*	Tajikistan	40.4	69.9	560
47	PI 606615	*L. culinaris subsp. culinaris*	Russia	45.8	38.6	-
48	ILWL 7	*L. culinaris subsp. orientalis*	Turkey	38.0	-	-

Appendix 6 Effect of *FTa2* 5' 2830-bp deletion on flowering time



(A) F_2 population derived from a cross between cv. Indianhead and ILL 5588 and evaluated under LD. (B) Mean *days to flowering* (DTF) for F_2 progeny grouped according to their *FTa2* 5' haplotype. (C) Mean *node of flower development* (NFD) for F_2 progeny grouped according to their *FTa2* 5' haplotype. Plants were evaluated under 12-h natural light extended with 4-h fluorescent light (LD). Data are mean ±SE for *n*=22-27.

Appendix 7 Co-segregation of ILL 5588 x ILL 2601 F_3 progeny with *MYB1*.

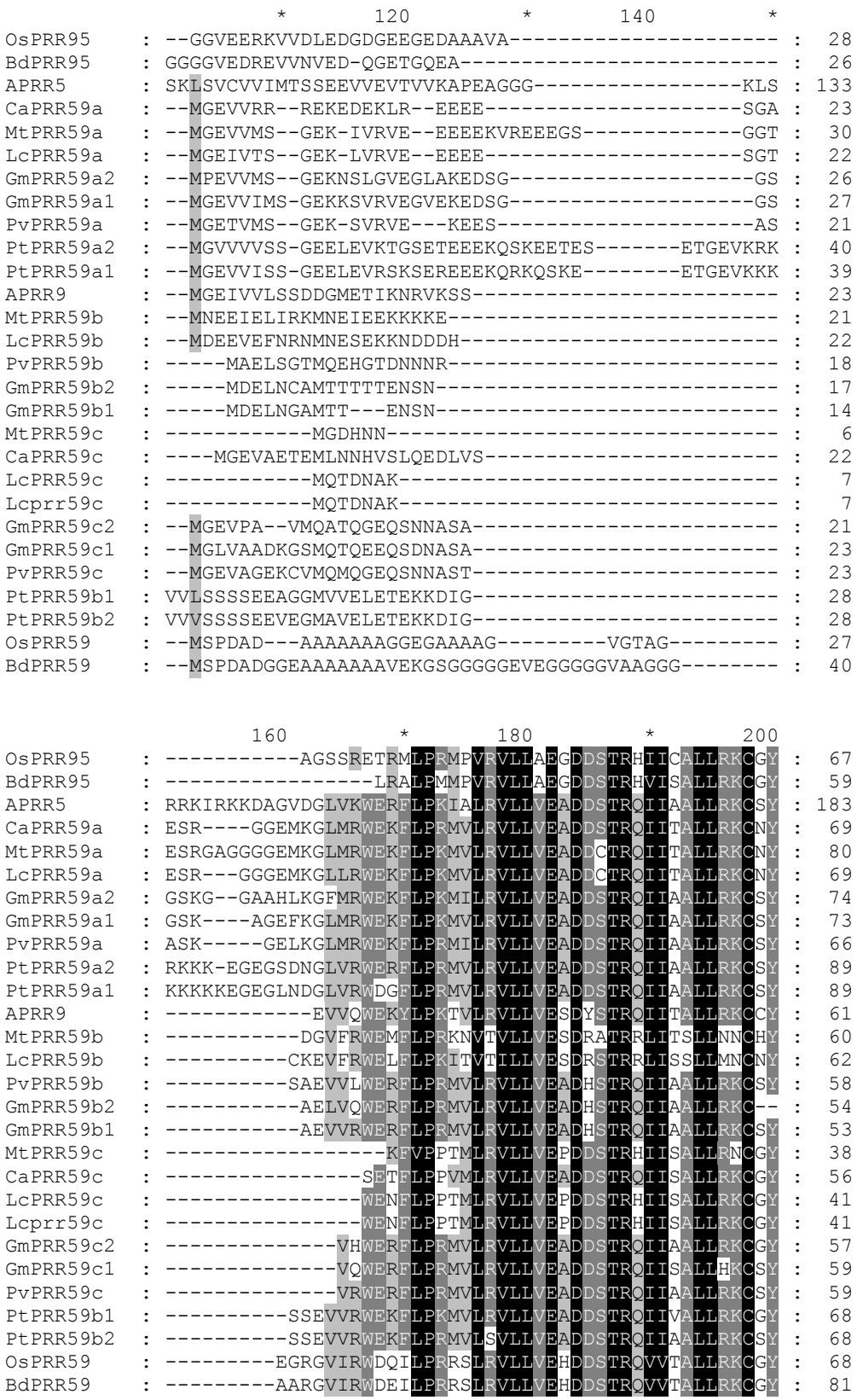


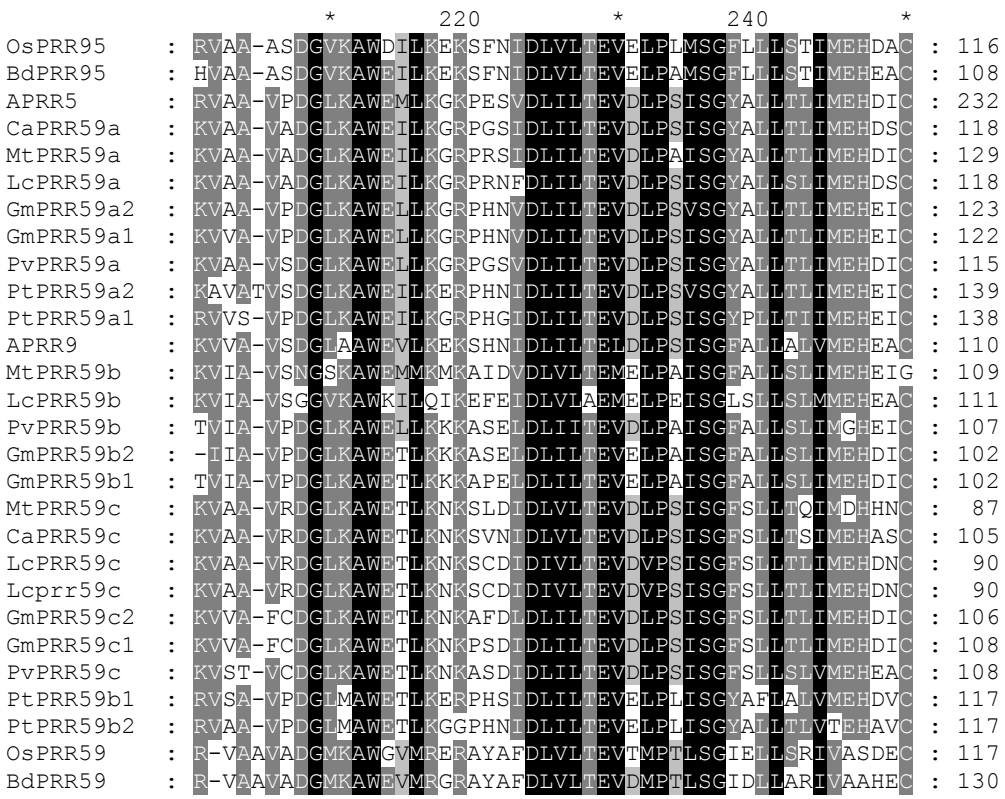
(A) ILL 2601 x ILL 5588 plant 163 F_3 progeny evaluated under SD conditions for flowering time. (B) Mean DTF for progeny for each allele in the ILL 2601 x ILL 5588 plant 163 F_3 progeny under SD conditions. Data are ±SE for *n*=3-18.

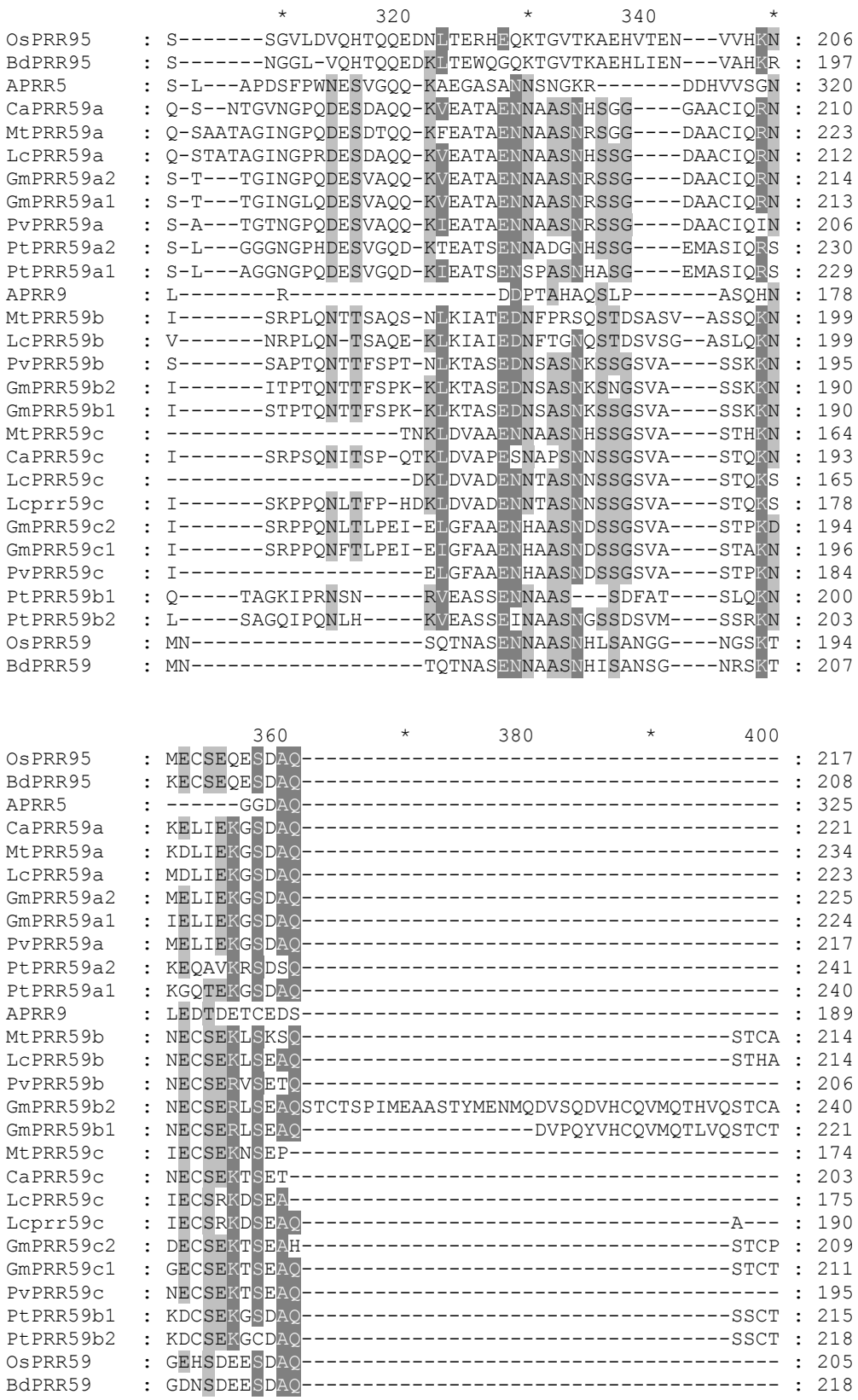
Appendix 8 *PRR5/9 full-length predicted protein alignment*

```
                    *        20         *        40         *
OsPRR95     : ---------------------------------------------- :   -
BdPRR95     : ---------------------------------------------- :   -
APRR5       : MWQTWPRQPILLDIFSNPNTLSTTVRSWSVRHPLSIITVKTFARFFLDIF :  50
CaPRR59a    : ---------------------------------------------- :   -
MtPRR59a    : ---------------------------------------------- :   -
LcPRR59a    : ---------------------------------------------- :   -
GmPRR59a2   : ---------------------------------------------- :   -
GmPRR59a1   : ---------------------------------------------- :   -
PvPRR59a    : ---------------------------------------------- :   -
PtPRR59a2   : ---------------------------------------------- :   -
PtPRR59a1   : ---------------------------------------------- :   -
APRR9       : ---------------------------------------------- :   -
MtPRR59b    : ---------------------------------------------- :   -
LcPRR59b    : ---------------------------------------------- :   -
PvPRR59b    : ---------------------------------------------- :   -
GmPRR59b2   : ---------------------------------------------- :   -
GmPRR59b1   : ---------------------------------------------- :   -
MtPRR59c    : ---------------------------------------------- :   -
CaPRR59c    : ---------------------------------------------- :   -
LcPRR59c    : ---------------------------------------------- :   -
Lcprr59c    : ---------------------------------------------- :   -
GmPRR59c2   : ---------------------------------------------- :   -
GmPRR59c1   : ---------------------------------------------- :   -
PvPRR59c    : ---------------------------------------------- :   -
PtPRR59b1   : ---------------------------------------------- :   -
PtPRR59b2   : ---------------------------------------------- :   -
OsPRR59     : ---------------------------------------------- :   -
BdPRR59     : ---------------------------------------------- :   -


                   60         *        80         *       100
OsPRR95     : ----------------------------------------------MG- :   2
BdPRR95     : ----------------------------------------------MGR :   3
APRR5       : FSPHYYRKNKVLFFALFSFISPLTNILICFVTVSLSLELSSSSSIIDLGF : 100
CaPRR59a    : ---------------------------------------------- :   -
MtPRR59a    : ---------------------------------------------- :   -
LcPRR59a    : ---------------------------------------------- :   -
GmPRR59a2   : ---------------------------------------------- :   -
GmPRR59a1   : ---------------------------------------------- :   -
PvPRR59a    : ---------------------------------------------- :   -
PtPRR59a2   : ---------------------------------------------- :   -
PtPRR59a1   : ---------------------------------------------- :   -
APRR9       : ---------------------------------------------- :   -
MtPRR59b    : ---------------------------------------------- :   -
LcPRR59b    : ---------------------------------------------- :   -
PvPRR59b    : ---------------------------------------------- :   -
GmPRR59b2   : ---------------------------------------------- :   -
GmPRR59b1   : ---------------------------------------------- :   -
MtPRR59c    : ---------------------------------------------- :   -
CaPRR59c    : ---------------------------------------------- :   -
LcPRR59c    : ---------------------------------------------- :   -
Lcprr59c    : ---------------------------------------------- :   -
GmPRR59c2   : ---------------------------------------------- :   -
GmPRR59c1   : ---------------------------------------------- :   -
PvPRR59c    : ---------------------------------------------- :   -
PtPRR59b1   : ----------------------------------------------MGK :   3
PtPRR59b2   : ----------------------------------------------MGE :   3
OsPRR59     : ---------------------------------------------- :   -
BdPRR59     : ---------------------------------------------- :   -
```


OsPRR95 : --GGVEERKVVDLEDGDGEEGEDAAAVA---------- : 28
BdPRR95 : GGGGVEDREVVNVED-QGETGQEA---------- : 26
APRR5 : SKLSVCVVIMTSSEEVVEVTVVKAPEAGGG----------KLS : 133
CaPRR59a : --MGEVVRR--REKEDEKLR--EEEE----------SGA : 23
MtPRR59a : --MGEVVMS--GEK-IVRVE--EEEEKVREEEGS----------GGT : 30
LcPRR59a : --MGEIVTS--GEK-LVRVE--EEEE----------SGT : 22
GmPRR59a2 : --MPEVVMS--GEKNSLGVEGLAKEDSG----------GS : 26
GmPRR59a1 : --MGEVVIMS-GEKKSVRVEGVEKEDSG----------GS : 27
PvPRR59a : --MGETVMS--GEK-SVRVE---KEES----------AS : 21
PtPRR59a2 : --MGVVVVSS-GEELEVKTGSETEEEKQSKEETES-------ETGEVKRK : 40
PtPRR59a1 : --MGEVVISS-GEELEVRSKSEREEEKQRKQSKE--------ETGEVKKK : 39
APRR9 : --MGEIVVLSSDDGMETIKNRVKSS---------- : 23
MtPRR59b : --MNEEIELIRKMNEIEEKKKKE---------- : 21
LcPRR59b : --MDEEVEFNRNMNESEKKNDDDH---------- : 22
PvPRR59b : -----MAELSGTMQEHGTDNNNR---------- : 18
GmPRR59b2 : -----MDELNCAMTTTTTENSN---------- : 17
GmPRR59b1 : -----MDELNGAMTT---ENSN---------- : 14
MtPRR59c : ------------MGDHNN---------- : 6
CaPRR59c : ----MGEVAETEMLNNHVSLQEDLVS---------- : 22
LcPRR59c : ------------MQTDNAK---------- : 7
Lcprr59c : ------------MQTDNAK---------- : 7
GmPRR59c2 : --MGEVPA--VMQATQGEQSNNASA---------- : 21
GmPRR59c1 : --MGLVAADKGSMQTQEEQSDNASA---------- : 23
PvPRR59c : --MGEVAGEKCVMQMQGEQSNNAST---------- : 23
PtPRR59b1 : VVLSSSSEEAGGMVVELETEKKDIG---------- : 28
PtPRR59b2 : VVVSSSSEEVEGMAVELETEKKDIG---------- : 28
OsPRR59 : --MSPDAD---AAAAAAAGGEGAAAAG----------VGTAG---------- : 27
BdPRR59 : --MSPDADGGEAAAAAAAVEKGSGGGGGEVEGGGGGVAAGGG-------- : 40
OsPRR95 : ----------AGSSRETRMLPRMPVRVLLAEGDDSTRHIICALLRKCGY : 67
BdPRR95 : ----------LRALPMMPVRVLLAEGDDSTRHVISALLRKCGY : 59
APRR5 : RRKIRKKDAGVDGLVKWERFLPKIALRVLLVEADDSTRQIIAALLRKCSY : 183
CaPRR59a : ESR----GGEMKGLMRWEKFLPRMVLRVLLVEADDSTRQIITALLRKCNY : 69
MtPRR59a : ESRGAGGGGEMKGLMRWEKFLPKMVLRVLLVEADDCTRQIITALLRKCNY : 80
LcPRR59a : ESR---GGGEMKGLLRWEKFLPKMVLRVLLVEADDCTRQIITALLRKCNY : 69
GmPRR59a2 : GSKG--GAAHLKGFMRWEKFLPKMILRVLLVEADDSTRQIIAALLRKCSY : 74
GmPRR59a1 : GSK----AGEFKGLMRWEKFLPKMVLRVLLVEADDSTRQIIAALLRKCSY : 73
PvPRR59a : ASK-----GELKGLMRWEKFLPRMILRVLLVEADDSTRQIIAALLRKCSY : 66
PtPRR59a2 : RKKK-EGEGSDNGLVRWERFLPRMVLRVLLVEADDSTRQIIAALLRKCSY : 89
PtPRR59a1 : KKKKKEGEGLNDGLVRWDGFLPRMVLRVLLVEADDSTRQIIAALLRKCSY : 89
APRR9 : ------------EVVQWEKYLPKTVLRVLLVESDYSTRQIITALLRKCCY : 61
MtPRR59b : ----------DGVFRWEMFLPRKNVTVLLVESDRATRRLITSLLNNCHY : 60
LcPRR59b : ----------CKEVFRWELFLPKITVTILLVESDRSTRRLISSLLMNCNY : 62
PvPRR59b : ----------SAEVVLWERFLPRMVLRVLLVEADHSTRQIIAALLRKCSY : 58
GmPRR59b2 : ----------AELVQWERFLPRMVLRVLLVEADHSTRQIIAALLRKC-- : 54
GmPRR59b1 : ----------AEVVRWERFLPRMVLRVLLVEADHSTRQIIAALLRKCSY : 53
MtPRR59c : ------------------KFVPPTMLRVLLVEPDDSTRHIISALLRNCGY : 38
CaPRR59c : ----------------SETFLPPVMLRVLLVEADDSTRQIISALLRKCGY : 56
LcPRR59c : ----------------WENFLPPTMLRVLLVEPDDSTRHIISALLRKCGY : 41
Lcprr59c : ----------------WENFLPPTMLRVLLVEPDDSTRHIISALLRKCGY : 41
GmPRR59c2 : --------------VHWERFLPRMVLRVLLVEADDSTRQIIAALLRKCGY : 57
GmPRR59c1 : --------------VQWERFLPRMVLRVLLVEADDSTRQIISALLHKCSY : 59
PvPRR59c : --------------VRWERFLPRMVLRVLLVEADDSTRQIIAALLRKCSY : 59
PtPRR59b1 : ----------SSEVVRWEKFLPKMVLRVLLVEADDSTRQIIVALLRKCGY : 68
PtPRR59b2 : ----------SSEVVRWEKFLPRMVLSVLLVEADDSTRQIIAALLRKCSY : 68
OsPRR59 : ---------EGRGVIRWDQILPRRSLRVLLVEHDDSTRQVVTALLRKCGY : 68
BdPRR59 : ---------AARGVIRWDEILPRRSLRVLLVEHDDSTRQVVTALLRKCGY : 81


* 220 * 240 *
OsPRR95 : RVAA-ASDGVKAWDILKEKSFNIDLVLTEVELPLMSGFLLLSTIMEHDAC : 116
BdPRR95 : HVAA-ASDGVKAWEILKEKSFNIDLVLTEVELPAMSGFLLLSTIMEHEAC : 108
APRR5 : RVAA-VPDGLKAWEMLKGKPESVDLILTEVDLPSISGYALLTLIMEHDIC : 232
CaPRR59a : KVAA-VADGLKAWEILKGRPGSIDLILTEVDLPSISGYALLTLIMEHDSC : 118
MtPRR59a : KVAA-VADGLKAWEILKGRPRSIDLILTEVDLPAISGYALLTLIMEHDIC : 129
LcPRR59a : KVAA-VADGLKAWEILKGRPRNFDLILTEVDLPSISGYALLSLIMEHDSC : 118
GmPRR59a2 : KVAA-VPDGLKAWELLKGRPHNVDLILTEVDLPSVSGYALLTLIMEHEIC : 123
GmPRR59a1 : KVVA-VPDGLKAWELLKGRPHNVDLILTEVDLPSISGYALLTLIMEHEIC : 122
PvPRR59a : KVAA-VSDGLKAWELLKGRPGSVDLILTEVDLPSISGYALLTLIMEHDIC : 115
PtPRR59a2 : KAVATVSDGLKAWEILKERPHNIDLILTEVDLPSVSGYALLTLIMEHEIC : 139
PtPRR59a1 : RVVS-VPDGLKAWEILKGRPHGIDLILTEVDLPSISGYPLLTIIMEHEIC : 138
APRR9 : KVVA-VSDGLAAWEVLKEKSHNIDLILTELDLPSISGFALLALVMEHEAC : 110
MtPRR59b : KVIA-VSNGSKAWEMMKMKAIDVDLVLTEMELPAISGFALLSLIMEHEIG : 109
LcPRR59b : KVIA-VSGGVKAWKILQIKEFEIDLVLAEMELPEISGLSLLSLMMEHEAC : 111
PvPRR59b : TVIA-VPDGLKAWELLKKKASELDLIITEVDLPAISGFALLSLIMGHEIC : 107
GmPRR59b2 : -IIA-VPDGLKAWETLKKKASELDLILTEVELPAISGFALLSLIMEHDIC : 102
GmPRR59b1 : TVIA-VPDGLKAWETLKKKAPELDLILTEVELPAISGFALLSLIMEHDIC : 102
MtPRR59c : KVAA-VRDGLKAWETLKNKSLDIDLVLTEVDLPSISGFSLLTQIMDHHNC : 87
CaPRR59c : KVAA-VRDGLKAWETLKNKSVNIDLVLTEVDLPSISGFSLLTSIMEHASC : 105
LcPRR59c : KVAA-VRDGLKAWETLKNKSCDIDIVLTEVDVPSISGFSLLTLIMEHDNC : 90
Lcprr59c : KVAA-VRDGLKAWETLKNKSCDIDIVLTEVDVPSISGFSLLTLIMEHDNC : 90
GmPRR59c2 : KVVA-FCDGLKAWETLKNKAFDLDLILTEVDLPSISGFSLLTLIMEHDIC : 106
GmPRR59c1 : KVVA-FCDGLKAWETLKNKPSDIDLILTEVDLPSISGFSLLTLIMEHDIC : 108
PvPRR59c : KVST-VCDGLKAWETLKNKASDIDLILTEVDLPSISGFSLLSLVMEHEAC : 108
PtPRR59b1 : RVSA-VPDGLMAWETLKERPHSIDLILTEVELPLISGYAFLALVMEHDVC : 117
PtPRR59b2 : RVAA-VPDGLMAWETLKGGPHNIDLILTEVELPLISGYALLTLVTEHAVC : 117
OsPRR59 : R-VAAVADGMKAWGVMREFAYAFDLVLTEVTMPTLSGIELLSRIVASDEC : 117
BdPRR59 : R-VAAVADGMKAWEVMRGFAYAFDLVLTEVDMPTLSGIDLLARIVAAHEC : 130
260 * 280 * 300
OsPRR95 : KNIPVIMMSSNDSVSMVFKCMLKGAADFLVKPIRKNELRNLWQHVWRKQL : 166
BdPRR95 : KNIPVIMMSSNDAVSMVFKCMLKGAADFLVKPIRKNELRNLWQHVWRKQL : 158
APRR5 : KNIPVIMMSTQDSVNTVYKCMLKGAADYLVKPLRRNELRNLWQHVWRRQT : 282
CaPRR59a : KSIPVIMMSSQDSVSTVYKCMLRGAADYLVKPIRINELRNLWQHVWRRQS : 168
MtPRR59a : KSIPVIMMSSQDSVSTVYKCMLRGAADYLVKPIRINELRNLWQHVWRRQT : 179
LcPRR59a : KTIPVIMMSSQDSVSTVYKCMCRGAADYLVKPIRINELRNLWQHVWRRQS : 168
GmPRR59a2 : KNIPVIMMSSQDSISTVYKCMLRGAADYLVKPIRKNELRNLWQHVWRRQS : 173
GmPRR59a1 : KNIPVIMMSSQDSISTVYKCMLRGAADYLVKPIRKNELRNLWQHVWRRQS : 172
PvPRR59a : KNIPVIMMSSKDSISTVYKCMLRGAADYLVKPIRKNELRNLWQHVWRRQS : 165
PtPRR59a2 : KNIPVIMMSSQDSIKTVYKCMLRGAADYLVKPIRKNELRNLWQHVWRKQS : 189
PtPRR59a1 : KNIPVIMMSSQDSISTVYKCMLRGAADYLVKPLRKNELRNLWQHVWRRQS : 188
APRR9 : KNIPVIMMSSQDSIKMVLKCMLRGAADYLIKPMRKNELKNLWQHVWRRLT : 160
MtPRR59b : RNIPVIMMSSRDSRSTVMKCMCRGAADFLIKPVRKNELTNLWQHVWRKHV : 159
LcPRR59b : KDIPLIMSS-HDSRGMVMNCMCKGAADFLIKPVRKNELTNLWQHVWRKHV : 160
PvPRR59b : KNIPVIMMSSHDSVSMVLKCMLKGAVDFLIKPIRRNELVNLWQHVWRRHA : 157
GmPRR59b2 : KSIPVIMMSSHDSVNMALKCMLNGAVDFLIKPIRKNELRNLWQHVWRRHT : 152
GmPRR59b1 : KNIPVIMMSSHDSVSMALKCMLKGAVDFLIKPIRKNELRNLWQHVWRRHA : 152
MtPRR59c : KNIPLIMMSSQDSVSTVFKFMLNGAVDFLIKPVRRNELRNLWQHVWRRNT : 137
CaPRR59c : KNIPVIMMSSHDSVSTAFKCMLNGAVDFLIKPVRRNELRNLWQHVWRRHT : 155
LcPRR59c : KKIPVIMSS-HDSVNTVFKCMQNGAVDFLIKPVRRNELRNLWQHVWRRHT : 139
Lcprr59c : KKIPVIMMSSHDSVNTVFKCMQNGAVDFLIKPVRRNELRNLWQHVWRRHT : 140
GmPRR59c2 : KNIPVIMMSSHDSVSMVFKCMLKGAADFLIKPVRKNELRNLWQHVWRRHA : 156
GmPRR59c1 : KNIPVIMMSSHDSVSMVLKCMLKGAADFLIKPVRRNELRNLWQHVWRRHA : 158
PvPRR59c : KNIPVIMMSSHDSVSMAFRCMLKGAADFLIKPVRKNELRNLWQHVWRWHV : 158
PtPRR59b1 : KNIPVIMMSSHDSISVVLKCMLKGSADFLVKPVRKNELRNLWQHVWRRQT : 167
PtPRR59b2 : KNIPVIMMSSQDSISMVLKCMLKGAADFLIKPVRKNELRNLWQHVWRRQT : 167
OsPRR59 : KNIPVIMMSSQDSIGTVLRCMQKGAVDFLVKPVRKNELRNLWQHVWRRHA : 167
BdPRR59 : KNIPVIMMSSQDSIGTVLRCMQNGAVDFLVKPVRKNELRNLWQHVWRRHS : 180


OsPRR95 : S-------SGVLDVQHTQQEDNLTERHEQKTGVTKAEHVTEN---VVHKN : 206
BdPRR95 : S-------NGGL-VQHTQQEDKLTEWQGQKTGVTKAEHLIEN---VAHKR : 197
APRR5 : S-L---APDSFPWNESVGQQ-KAEGASANNSNGKR-------DDHVVSGN : 320
CaPRR59a : Q-S--NTGVNGPQDESDAQQ-KVEATAENNAASNHSGG----GAACIQRN : 210
MtPRR59a : Q-SAATAGINGPQDESDTQQ-KFEATAENNAASNRSGG----DAACIQRN : 223
LcPRR59a : Q-STATAGINGPRDESDAQQ-KVEATAENNAASNHSSG----DAACIQRN : 212
GmPRR59a2 : S-T---TGINGPQDESVAQQ-KVEATAENNAASNRSSG----DAACIQRN : 214
GmPRR59a1 : S-T---TGINGLQDESVAQQ-KVEATAENNAASNRSSG----DAACIQRN : 213
PvPRR59a : S-A---TGTNGPQDESVAQQ-KIEATAENNAASNRSSG----DAACIQIN : 206
PtPRR59a2 : S-L---GGGNGPHDESVGQD-KTEATSENNADGNHSSG----EMASIQRS : 230
PtPRR59a1 : S-L---AGGNGPQDESVGQD-KIEATSENSPASNHASG----EMASIQRS : 229
APRR9 : L--------R----------DDPTAHAQSLP-------ASQHN : 178
MtPRR59b : I-------SRPLQNTTSAQS-NLKIATEDNFPRSQSTDSASV--ASSQKN : 199
LcPRR59b : V-------NRPLQN-TSAQE-KLKIAIEDNFTGNQSTDSVSG--ASLQKN : 199
PvPRR59b : S-------SAPTQNTTFSPT-NLKTASEDNSASNKSSGSVA----SSKKN : 195
GmPRR59b2 : I-------ITPTQNTTFSPK-KLKTASEDNSASNKSNGSVA----SSKKN : 190
GmPRR59b1 : I-------STPTQNTTFSPK-KLKTASEDNSASNKSSGSVA----SSKKN : 190
MtPRR59c : ----------TNKLDVAAENNAASNHSSGSVA----STHKN : 164
CaPRR59c : I-------SRPSQNITSP-QTKLDVAPESNAPSNNSSGSVA----STQKN : 193
LcPRR59c : ----------DKLDVADENNTASNNSSGSVA----STQKS : 165
Lcprr59c : I-------SKPPQNLTFP-HDKLDVADENNTASNNSSGSVA----STQKS : 178
GmPRR59c2 : I-------SRPPQNLTLPEI-ELGFAAENHAASNDSSGSVA----STPKD : 194
GmPRR59c1 : I-------SRPPQNFTLPEI-ELGFAAENHAASNDSSGSVA----STAKN : 196
PvPRR59c : I----------ELGFAAENHAASNDSSGSVA----STPKN : 184
PtPRR59b1 : Q-----TAGKIPRNSN-----RVEASSENNAAS---SDFAT----SLQKN : 200
PtPRR59b2 : L-----SAGQIPQNLH-----KVEASSEINAASNGSSDSVM----SSRKN : 203
OsPRR59 : MN----------SQTNASENNAASNHLSANGG----NGSKT : 194
BdPRR59 : MN----------TQTNASENNAASNHISANSG----NRSKT : 207
OsPRR95 : MECSEQESDAQ---------- : 217
BdPRR95 : KECSEQESDAQ---------- : 208
APRR5 : ------GGDAQ---------- : 325
CaPRR59a : KELIEKGSDAQ---------- : 221
MtPRR59a : KDLIEKGSDAQ---------- : 234
LcPRR59a : MDLIEKGSDAQ---------- : 223
GmPRR59a2 : MELIEKGSDAQ---------- : 225
GmPRR59a1 : IELIEKGSDAQ---------- : 224
PvPRR59a : MELIEKGSDAQ---------- : 217
PtPRR59a2 : KEQAVKRSDSQ---------- : 241
PtPRR59a1 : KGQTEKGSDAQ---------- : 240
APRR9 : LEDTDETCEDS---------- : 189
MtPRR59b : NECSEKLSKSQ----------STCA : 214
LcPRR59b : NECSEKLSEAQ----------STHA : 214
PvPRR59b : NECSERVSETQ---------- : 206
GmPRR59b2 : NECSERLSEAQSTCTSPIMEAASTYMENMQDVSQDVHCQVMQTHVQSTCA : 240
GmPRR59b1 : NECSERLSEAQ----------DVPQYVHCQVMQTLVQSTCT : 221
MtPRR59c : IECSEKNSEP---------- : 174
CaPRR59c : NECSEKTSET---------- : 203
LcPRR59c : IECSRKDSEA---------- : 175
Lcprr59c : IECSRKDSEAQ----------A--- : 190
GmPRR59c2 : DECSEKTSEAH----------STCP : 209
GmPRR59c1 : GECSEKTSEAQ----------STCT : 211
PvPRR59c : NECSEKTSEAQ---------- : 195
PtPRR59b1 : KDCSEKGSDAQ----------SSCT : 215
PtPRR59b2 : KDCSEKGCDAQ----------SSCT : 218
OsPRR59 : GEHSDEESDAQ---------- : 205
BdPRR59 : GDNSDEESDAQ---------- : 218

```
                    *       420         *       440         *
OsPRR95     : ----------------------------SSCTRSELEADSRQTNN----L : 235
BdPRR95     : ----------------------------SSCTRSEVEAESKHTNN----F : 226
APRR5       : ----------------------------SSCTRPEMEGESADVEVS---A : 344
CaPRR59a    : ----------------------------SSCTKPNMEAESGSLVD---IA : 240
MtPRR59a    : ----------------------------SSCTRPNMEAESG-LVD---NM : 252
LcPRR59a    : ----------------------------SSCTKPNMEMESG-LVD---NM : 241
GmPRR59a2   : ----------------------------SSCTKPDCEAESG-PVDNIDNI : 246
GmPRR59a1   : ----------------------------SSCTKPDCEAESD-PVG---NM : 242
PvPRR59a    : ----------------------------SSCTKPDLEAESG-PVD---NT : 235
PtPRR59a2   : ----------------------------SSCTKPGLEAEGAHMEN----M : 259
PtPRR59a1   : ----------------------------SSCTKPDLEAESSHMEN----M : 258
APRR9       : --------------------------------RYHSDQGSGAQ----A : 201
MtPRR59b    : MPFSDAKNLYMDNMQKPCQMKS--SVKLRNIDVLKHAESNKIERG----S : 258
LcPRR59b    : LPFSEAENACTDNMQNASQMKS--CFKLSKIGGLKHKESNQLERE----S : 258
PvPRR59b    : ---------------DMPQLK---NSKQKKIDLVKHEKFARFESE----S : 234
GmPRR59b2   : SPIFEAESTFVENMQDVPQLE---SSKLNKIDMVDHEKFAKFERK----S : 283
GmPRR59b1   : SPIFEAKSTYVENMQDVPPLK---SSKLNKIDMVKHEKFAQFERE----S : 264
MtPRR59c    : --------------------------------HDNSIKYERE----S : 185
CaPRR59c    : --------------QDMSQLKS--SLSLSNTDKVKQESSIKFERE----S : 233
LcPRR59c    : -----------------------------------KKSTKYEWE----S : 185
Lcprr59c    : --------------QDMSQLKS--SSSLSNTDQVKHKKSTKYEWE----S : 220
GmPRR59c2   : SPFLEAESTYMENMQDILQLKS--SSNLSNIDTVKHENSTKCERE----S : 253
GmPRR59c1   : SPFLEAESTYLENMQDISQLKR--SSNLSNIDTVKHENSTKCERE----S : 255
PvPRR59c    : ---------------DISLLKS--SSNLSDIGTVKHENSTKCERE----S : 224
PtPRR59b1   : TPCLEAESAHMQNIQGLSYLKYRSASNLSDADNEKYEDYAKLNKSPVNPE : 265
PtPRR59b2   : TPCLEAESAHMQNMQGLSQMKYRSASNLSNTDREEFEECAKLDKSPVTPE : 268
OsPRR59     : ---------------SSGSKREVEIQSAEKLPEVVADGGAGSSREHKIQN : 240
BdPRR59     : ---------------SSGSKRETEIQSVEKLPETVTENGASSSRELTIQN : 253

                   460         *       480         *       500
OsPRR95     : LEYKQPMGRHFSKPDHKNTEKNGGTKIHASNDGNLIPRREEDASLR---- : 281
BdPRR95     : LEFKQITGKYLS-SDLKSTEDNGQTKTQTIRDDNLIPRRERDLSPR---- : 271
APRR5       : RDAVQMECAKS--------------QFNETRLLAN---------------- : 365
CaPRR59a    : QEFSPTKCAEAYPSGIQTHEVD--IQLGQASTPQDNHDRGLSVANC---- : 284
MtPRR59a    : HEFSQLKCAEAYPSEIKTRELD--IHLGQAVIAQDSHAGGLSVANCNNGV : 300
LcPRR59a    : QEFTQLKCAEACPSEIKTQEFD--IRLGQTLITQDSHAGGLRTANR---- : 285
GmPRR59a2   : QEFSPLKCGEAYPSGTETQQVETSIRLGQTLVMHASHAGGLNVSIC---- : 292
GmPRR59a1   : QEFSLLKCGEAYPSGTETQQVETSFRLGQTLVMHDCHAGGLNVSIR---- : 288
PvPRR59a    : LEFSPAKCGEEYPNGAETQEVETCIRLGRTLVMNDSHAGGL--TMQ---- : 279
PtPRR59a2   : QEFLQPVWSKFSLTDTNMQKHEEHVNLGQKLLVRDSEAEGSATAVC---- : 305
PtPRR59a1   : QEFLQPVRSIFSLTDMNMQKREMHVNLGQKLLLHDREAEGSAAAAR---- : 304
APRR9       : INYN------------------------------------------- : 205
MtPRR59b    : TKQNDETGDS--------------------RLEQDCSTAE--IEPK---- : 282
LcPRR59b    : SKLNDEARGS--------------------RLEQDYSTSE--VEPK---- : 282
PvPRR59b    : AKINNETRDKSITIVSDTARCDKTFESTDLRLEQDHCCAD--TEIE---- : 278
GmPRR59b2   : AKHNDETEDKSITIVSEAARCDKSFELTDLMLEQDCGVAEPETENE---- : 329
GmPRR59b1   : AEHNDETEDKSVTIVSDAARCDKTSELTELRPEQDCGVAEPETENE---- : 310
MtPRR59c    : AEYKDVTGEKSTTIASKAAGCDKIS--TGLRLGQNYDYSETENRD----- : 228
CaPRR59c    : AEYNNETGEKSTTLVSKVARYDKIS--TGLRLGQCYDHSETENQDQ---- : 277
LcPRR59c    : AEYNGETGEKSTSTAPKAAGRDKIS--TGLRLGQSYEYNETENRD----- : 228
Lcprr59c    : AEYNGETGEKSTSTAPKAAGRDKIS--TGLRLGQSYEYNETENRD----- : 263
GmPRR59c2   : DKHNDEAGEKSLFILEDARCNKTFKP-TGLRLGQSYECHETRNQD----- : 297
GmPRR59c1   : DKHNDEAGEKSLFVSEDSRCNKTFKP-TGLRLGQGYECCETRNQD----- : 299
PvPRR59c    : DEHNDEAGAKSIVISEDEGCNKTFKP-TGLRLGQGYDFGEMGNQD----- : 268
PtPRR59b1   : SKTGVFVAERSNRTRPDREPYHGAYNPTASRLVEEHACAKSAIHD----- : 310
PtPRR59b2   : NKTGVFVPERPNRMESDGEPCSGAYNPTSLRLLEEHACAKSAIQD----- : 313
OsPRR59     : GFIDGMNTKSHALKGNDDAPSGNACGDSELQVLSTEKNVRSKFLNG---- : 286
BdPRR59     : GPFDRVTTKAHAFNINVDPPSGNVCETGELQVFSAEKKLRSKCLNG---- : 299
```

```
                    *       520         *       540         *
OsPRR95     : ----------------------------RMTCSNDINCEKASRDMELVH : 302
BdPRR95     : ----------------------------KQPCLKDNDCQKATREIEVVH : 292
APRR5       : ---------------------------------------------ELQSKQAE : 373
CaPRR59a    : --------KNGETNTNNDKDGDD------PSISGEVHDNHYDQTDS-SKE : 319
MtPRR59a    : ASTN--NCKNGETGANNCKDGDDQEHFRNASISGEVHDNHYVQIYSTTKE : 348
LcPRR59a    : --------KNKETSTNNGKSVDDQEHFRTASISGEVHDNHYVQINS-TKE : 326
GmPRR59a2   : --------KNGEASTT---DADP-EHFGNG-ISGEAHDNHYVQMNS-SKE : 328
GmPRR59a1   : --------KNGEASTTNDKDTDT-EHFGNASISGEAHDNPYVQINS-SKE : 328
PvPRR59a    : --------RNGEASTTNDKDADP-EHLGNASISGEAHDNHYVQINS-SRE : 319
PtPRR59a2   : ---------EDSNKITVDKEITPGSGRVTANIAIEGCDKIGALANS-PRE : 345
PtPRR59a1   : ---------EDANIMDVDKEISPGNGRTGAYVAIESCDNDVALANS-HRE : 344
APRR9       : ------------------------GHNKLMENGKSVDERDEFKET-FDV : 229
MtPRR59b    : -----------------CEIFKAESSRENPDIDTEIRECSNELIEP-SSR : 314
LcPRR59b    : -----------------NEIFRAELSRENPDTDSEIRGCSDGLMEP-SSK : 314
PvPRR59b    : -----------------VEILKYDLGIGDSNISTELHEWSDERVKP-IKG : 310
GmPRR59b2   : -----------------DEILKSELGRDNSHVS-ILHGCNAEQVKP-SKG : 360
GmPRR59b1   : -----------------DEILKSELDGDNSHVS-MMQGCSAERVKP-SKG : 341
MtPRR59c    : -----------------EVLGTELSKAHPHINTKIHQSNNELEDH-SAG : 259
CaPRR59c    : -----------------DEVLRTELSKADPHINRKIHRCNDELVEH-CTG : 309
LcPRR59c    : -----------------EGLKTELGKANPRVNTKIRERNDEKEEH-SAG : 259
Lcprr59c    : -----------------EGLKTELGKANPHVNTKIRERNDEKEEH-SAG : 294
GmPRR59c2   : -----------------EVLRIELIKSNPEINTDIHRCSDELVDP-STG : 328
GmPRR59c1   : -----------------VVLKIELIKSNPEINTDIHGCSDELVDP-YTG : 330
PvPRR59c    : -----------------EVLRIELSKANPEINVDIHGCIDELEGP-STG : 299
PtPRR59b1   : ---------------------ENSRPENDREHANSSFGHDDVLAET-SSG : 338
PtPRR59b2   : ---------------------ENSRPENDRGLANSSFGCDDVPFES-SSG : 341
OsPRR59     : ----ITSAKVAGQIMDNALRFADSSSLRSSDPGKDLLVVAQTTADRKCKS : 332
BdPRR59     : ----ITSAKVAGQIMDNALRIADASSCRPTDPGKDLLAAAPSTAGKKGNS : 345

                   560         *       580         *       600
OsPRR95     : IIDNQQKNNTHMEMDVARANSRGNDDKCTSIPAHQLELSLRR----SDYS : 348
BdPRR95     : IIDDEQKSNAQTDVDVMRTTFHGNCDKGTSIPAHQLELSLRR----SDYS : 338
APRR5       : AIDFMGASFRRTG---RRNRE--ESVAQMES-RIELDLSLRR----PNAS : 413
CaPRR59a    : AIDLIGAFRSRPNCGLNNSST--DCTGKTEY-SPQLDLSLRS----SHHS : 362
MtPRR59a    : AIDLIGAFRTDPNCSLKNSSI--DCTGKTDH-SPQLDLSLRS----SHPS : 391
LcPRR59a    : AIDLIGAFRTHPNCSLKKSSI--DWTDKLDN-SPQLNLSLRS----SHPS : 369
GmPRR59a2   : AIDFIGAFHTHPICTLKNSTV--NCTGKTDL-SPQLDLSLRR----SRPS : 371
GmPRR59a1   : AMDLIGAFHTHPNCSLKNSTV--NCTGNTDH-SPQLDLSLRR----SCPG : 371
PvPRR59a    : AIDLIGAFHTHPNCMVKSPTV--DCTGKVDL-SPQLDLSLRR----SRPS : 362
PtPRR59a2   : AIDFMGASTN------HSSFN--NVEIHTCS-SPHLDLSLRR----SHPS : 382
PtPRR59a1   : AFDFMGASTNR-----SSSFN--NVKINTDS-SPHLDLSLRR----SHPS : 382
APRR9       : TMDLIGGIDKRPD------SIYKDKSRDECV-GPELGLSLKR-SCSVSFE : 271
MtPRR59b    : AVDLISTFGNLHKRTKEIHVTNGDKETKTDF-EKELELSLRS-DFSGSSC : 362
LcPRR59b    : AIDLISTVGNLRKCTKEIHWIKGDKETKTEL-EKELERSLRS-DFSGSSC : 362
PvPRR59b    : AIDLIATFGNLPKHPDENCNLNGGNTTMTDG-VTQLELSLRS-DFPGSSC : 358
GmPRR59b2   : AIDLIATFGNLPKHPNENCSLNGGNTTKTDC-ETQLELSLRS-DFPGSSG : 408
GmPRR59b1   : AIDLIATVGNLPKHLDENCSLNGGNTTKTDC-ETQLELSLRS-DFPGSSG : 389
MtPRR59c    : AIDLMATFDKYPKNNHANCSFSGGNTAKTDF-DTQFELSLQR-DSPGSPP : 307
CaPRR59c    : AIDLIATIENLPKSSYANCSFNGGNTAKTDF-DSQLELSLQR-DFRGSSP : 357
LcPRR59c    : AIDLMATFENLPKSSYADCSFNDGNRAKTED-DTQLELSLQR-DYPGSSS : 307
Lcprr59c    : AIDLMATFENLPKSSYADCSFNDGNRAKTED-DTQLELSLQR-DYPGSSS : 342
GmPRR59c2   : AIDLIATFKNLPKSTDEKCSFSSGNTAKTDF-DTQLELSLRR-DFPGSSC : 376
GmPRR59c1   : AIDLIATFKNLPKSTDEKCSFSSGNTAKTDF-DTQLELSLRRRDFPGSSC : 379
PvPRR59c    : AIDLIGTFKTLPKSTDENCSFSSGNIAKTDF-DTQLELSVRR-YFPGSSY : 347
PtPRR59b1   : AIDLIGSFNNQPKHTYAYSSLHDA-TNKTEF-PPLLELSLRR-LYPSSSK : 385
PtPRR59b2   : AIDLIGTLNNGPKTTYVHSSLHYG-TNKTEF-APQLELSLKR-LYPSSSK : 388
OsPRR59     : SALEN---NAVMENNLSENSKGTATGHAESCPSHFVEINLEK-QHHLN-- : 376
BdPRR59     : PAIENSAVNPAMENTPHERSKGTAIGRAESCPPRSLEINLEK-QPLFNSN : 394
```

```
                      *       620         *       640         *
OsPRR95     : RLESQEKNERRTLNHSTSSPFSLYN---------------------- : 373
BdPRR95     : KLDDQEKNDKRTLNHSTSSAFSLYN---------------------- : 363
APRR5       : --ENQSSGDRPSLHPSSASAFTRYV--------------------HRPLQ : 441
CaPRR59a    : NFEKELNEERHTLMHSNASAFKRYT--------------------NRQLQ : 392
MtPRR59a    : NFEKDLTEERHTLMHSNASAFKRYT--------------------NRQLQ : 421
LcPRR59a    : NNEKELTEDRNTLMHSNASAFKRYT--------------------NRQLQ : 399
GmPRR59a2   : SFENELTEERHTLMHSNASAFKRYT--------------------NRQLQ : 401
GmPRR59a1   : SFENKLTEERHTLMHSNASAFKRYT--------------------TRQLQ : 401
PvPRR59a    : SFENELTEERHTLMHSNASAFKRYT--------------------NRQLQ : 392
PtPRR59a2   : GFETQVTEERHTLRHSNASAFTWYT--------------------NRASQ : 412
PtPRR59a1   : GFEIRDTEERRALWHSNASAFTQYI--------------------NRPLQ : 412
APRR9       : N---QDESKHQKLSLSDASAFSRFE------------------------- : 293
MtPRR59b    : KQASETTEEWQRLNHSNASAFSRYD--------------------GSKMLQ : 393
LcPRR59b    : KQASEATEEWQRLNHSNTSAFSRYD--------------------GSKMLW : 393
PvPRR59b    : KQASEATEESQRLNHSHNSAFSWYS--------------------NSKLVH : 389
GmPRR59b2   : KQASESTEESQRLNHSNTSAFSWYS--------------------NSKLLQ : 439
GmPRR59b1   : NQASEATEESQRLNHSNTSAFSWYS--------------------NSKLLQ : 420
MtPRR59c    : K---PTTEERQILNHSNASAFSWYG--------------------SSMVLQ : 335
CaPRR59c    : Q---PTTEERHILNHSNASAFSWYG--------------------SSKLLH : 385
LcPRR59c    : K---PTTEEGQILNHSNASPFSRYG--------------------SSMLLQ : 335
Lcprr59c    : K---PTTEEGQILNHSNASPFSRYG--------------------SSMLLQ : 370
GmPRR59c2   : K---AAFKERQILNHSNASAFSRYS--------------------NSKLLQ : 404
GmPRR59c1   : K---AASEERQLLNHSNASAFSRYS--------------------SSKLLQ : 407
PvPRR59c    : K---DVSEERQILNHSNASAFSRYG--------------------GSKLLQ : 375
PtPRR59b1   : N---QGLDERHALNHSNSSAFSL---------------------YNSKTLQ : 412
PtPRR59b2   : N---QGVDERHALNHSHASAFSWKKQGCWDSGRDGIGGSDFRRYNSKTLQ : 435
OsPRR59     : GYTNHKLNEKDIFNHSNSSAFSRYGNK----------------RIESSAQ : 410
BdPRR59     : GYANQEFKDKDNFRHSNSSAFSRYGNK----------------RIESSVQ : 428

                    660         *       680         *       700
OsPRR95     : ----------------CRTASSTINAGDAQACSTSATHIDLENKNGDSKTP : 408
BdPRR95     : ----------------CRTASSCGNAGDAHLCSTSATHVDLEMKTGDSVAP : 398
APRR5       : TQCSASPV-VTDQRKNVAASQDDNIVLMNQYNTSEPPPNAPRRNDTSFYT : 490
CaPRR59a    : AS-PAVVLNFPDQQREQKIN---NENHIAAGCNSDSS--IPSKQKCIVSP : 436
MtPRR59a    : AS-PAVVINFSDQPREQKTN---NENH-----NSDSS--IPSMQ------ : 454
LcPRR59a    : AS-HAVVVNFSDQQREQKTN---NDNH-----NSDSS--IPSKQKCNISP : 438
GmPRR59a2   : ISTPAVLINFSDQQRQQIANCEKNISRIATGCNSDSS--TPSMQRCIVSP : 449
GmPRR59a1   : ISMPAVLINFSDQQREQITNCEKNISHIATGSNSDSS--TP-MQRCIVSP : 448
PvPRR59a    : VLTPAALINFSDQQRDQRANGKESISHVVTGCNSDSS--TPGMQRYMVSP : 440
PtPRR59a2   : LPHSALAN--TGNQEEFRANYDGKISSNVNGYNSDALSLAPSTRRSAISL : 460
PtPRR59a1   : LPHSALES--TGNQKELGTNYDRKISS--TGYNSDALSLAPSTQKSEISL : 458
APRR9       : ------------------ESKSAEKAVVALEESTSGEPKTPTESHEK--- : 322
MtPRR59b    : QLLQNS----------NWNSNK-SQELSVVTAGNCFQYAGS-IKMEN--M : 429
LcPRR59b    : PLFQNS----------NWSSNK-SHELSVVTADNCIQYGGP-IKIED--M : 429
PvPRR59b    : PLFQTPSITSTEVNNPSWDSHE-SHKLSRTTSGNCCQYGGSNKNLEN--M : 436
GmPRR59b2   : PLFSPPSITSPKVNWLNWDSHE-CLKLS----GN-CQYDDSNQNLEN--M : 481
GmPRR59b1   : PHFSTPSITFPEVNNLSWDSHE-SHKLSGITSGN-CQYGGSNQNLEN--M : 466
MtPRR59c    : PLFP------------TKSSHE-SQKLS-ENINTTHHYDGKKQKQEN--I : 369
CaPRR59c    : PLFP------------SNNFHE-SHKLS-QDTNTT--------------- : 406
LcPRR59c    : PLFP------------TNNSYE-SQKLS-ENTNTTYECDGKNQKEES--N : 369
Lcprr59c    : PLFP------------TNNSYE-SQKLS-ENTNTTYECDGKNQKEES--N : 404
GmPRR59c2   : PLFPTPSTISAKLTNASQNSHE-SLKLS-KNTSTSHQYSEKSQNQEE-KI : 451
GmPRR59c1   : PLFPTPSTISAKLTNSSLSSHE-SHKFS-ENASTSHQYGGKNQNQE--KI : 453
PvPRR59c    : TLCPTPSTISAKLTNGSHDSQK-SHKLS-ENTSTSDQYGAENQIKE--KI : 421
PtPRR59b1   : SLFPTSASNGSDSKEEASKSPDPSSNQLAQNVGTLSQIHDASLSGNQEIM : 462
PtPRR59b2   : PPFPASASNGSDSKEEASKSPELSSNQHAQNINSISQRHGATLSGNQD-M : 484
OsPRR59     : RPFPPSFRVVHQQPVYDKNPQSSRVLLSCEHNTRESTVQAQVPLDRSTEG : 460
BdPRR59     : QLFPPSLHLSHHEPVCDKNIQPGGALSSREHNTWKSAVQAKVPLDSCTER : 478
```

```
                      *       720         *       740         *
OsPRR95     : -------SQDKRETNQPPIRVVPFPVPVGGLTFDGQPFWNGAPVASLFYP : 451
BdPRR95     : -------SQDKTDAICPPIRVVPFPVPVAGLTFNGQPFWSGAPVAPLLYP : 441
APRR5       : GADSPGPPFSNQLNSWPGQSSYPTPTPINNIQFRDPNTAYTSAMAP---- : 536
CaPRR59a    : --ATAQSKESELATSHS-QQGHSLPIPVKGVRFNDLCMAYGSALPPGFRT : 483
MtPRR59a    : -------KESELATSHS-QQRHSLPIPVKGVRFNDLCMAYGSTLPPGFRT : 496
LcPRR59a    : --ATAQSKESDLATSHS-QQGHSLPIPVKGLRFNDLCVAYGSTLPQGFCT : 485
GmPRR59a2   : --TTVQSKEPELATSHS-QPGHSLPIPVKGVRFNDLCTTYGSVFPSVFRA : 496
GmPRR59a1   : --TTVQSKESELATSHP-PQGHSLPIPVKGVRFNDLCTAYGSVLPSVFHT : 495
PvPRR59a    : --TTVQSKESELATSHS-QQVHSLPIPVKGVRFNDLWTAYGSVHPSMFRT : 487
PtPRR59a2   : --AAGQTKEYEIVTSSSGEKVFPIHIPVKDTRFNNLCNSYGAVLP----- : 503
PtPRR59a1   : --AAGQTKESEIATSSPGQRVFPIQIPAKETRLNNLCNSYGSVFPPIFCK : 506
APRR9       : ------------------------------------------------- :   -
MtPRR59b    : --TTAVMAQYEQLG---------------------------------LSA : 444
LcPRR59b    : --TNSVTAHYGQFGP-------------------------KLYNTGLFA : 451
PvPRR59b    : --ISTVIDQYGQVKPNLSNSQCGMLPVS-GVISDLKSKGHGNVLTSVFSA : 483
GmPRR59b2   : --ISTVIGQY------------GLLPVS-GVISKLKSEGHGHVFTSVFYA : 516
GmPRR59b1   : --IGTVICQYGQVTPKLSNSQCGLLPVS-GVISNLKSEGHGNVFTSLFYA : 513
MtPRR59c    : --TYLVIGQSGQVDTK---CQLEFFPAT-GATSDNKSMEHNNVLHSMFNA : 413
CaPRR59c    : --------QSGQVDAKLPNSLLEFFPAT-GSTFDNKSIGHGNVFHSMFYI : 447
LcPRR59c    : --TYLVIGQSGQVDAKFLNSQHEFFPATTGDSSDNKSMEHDNIFHSISNA : 417
Lcprr59c    : --TYLVIGQSGQVDAKFLNSQHEFFPATTGDSSDNKSMEHDNVFHSISNA : 452
GmPRR59c2   : --ITSVIGQSGQVDPKLPNSQLGLFPAT-GVTSDHKSKGNGNVFPSKLYA : 498
GmPRR59c1   : --ITPVIGQSGQVDPKLPNSQLGYFPAT-GVTSDHKSTGNGNVFPSMLYA : 500
PvPRR59c    : --TTSDIGQSEQVDPNLPNSQLRFFPAT-GVTYDHKSTGNGNVFPSMLYS : 468
PtPRR59b1   : --TTPVIGQSGKVELAHPSPQLGLIPVL-GTRLDNISTGCGHVFSPLCYT : 509
PtPRR59b2   : --TIPIIGQSGKAELAYPSPRHGLIPVR-RGMLDNISTEYGHDFSPLYYT : 531
OsPRR59     : ---AAILCSSSVREDAGTSSSSPRKDSLTHPSYGFIPVPIPVGAAIPYHY : 507
BdPRR59     : ---VAILSSSSAREDAGPSSSSPRTEILNHPPYGFIPVPIPVGAAIPYHY : 525

                    760         *       780         *       800
OsPRR95     : QSAPPIWNSKTSTWQDATTQAISLQQNGPKD-----TDTKQVENVEEQTA : 496
BdPRR95     : QSGPPIWNSKTSTSKQAAAQAILSQQKWQQSNATV-MDSDQAEITQGQEV : 490
APRR5       : ---ASLSPSPSSVSPHEYSSMFHPFNSKP-------------------EGL : 565
CaPRR59a    : QSGPPSTPG--SVTILEPNFQAEAFYQSNGKEN---NAEQPYEPRHSNGN : 528
MtPRR59a    : QSGPPSMPG--SVVFLEQNFQADAFYQSNVKQN---NSEQLYEPRGPNGN : 541
LcPRR59a    : QSGPPSMPG--SVVFLEQNFQADAFYQPNVKEN---NSEQLYEPRCTNGN : 530
GmPRR59a2   : QSGSPAMPSPNSVMLLEPNFQVNAFYQSNMKES---SSEQLYEPGGPNGN : 543
GmPRR59a1   : QSGPPAMPSPNSVVLLEPNFQVNAFYQSNMKES---SSEQLYESRGPNGN : 542
PvPRR59a    : QSGPPTMPSPSSVVLLEPNFQN-AFHQSNMKES---SSEQLYESLGPNGN : 533
PtPRR59a2   : -----PMMSQSSASQKEPIHKVNPFQCSN-YGS---TSVQLCDRLGQNAN : 544
PtPRR59a1   : QSGLSPMMSPSSACQQEPTYKVNQFQHSN-HGS---TSEQ--NRLGQHTN : 550
APRR9       : --------------LRKVTSDQGSATTSSNQEN---IGSSS--------- : 346
MtPRR59b    : DNVFHHMLTPKSNCQKESSPFPSSSSSQSNPES---HNSEHDHNCCYDAN : 491
LcPRR59b    : DNVLHRMWIPKSNFEKESSPFPSSSSSQSNPES---HNSGHHHNCSYDAN : 498
PvPRR59b    : PCGTHPVWSSKPVCQNESSSFPISTSSQSNPES---HNSDQYHDCSNIA- : 529
GmPRR59b2   : QSGIHPMLSPKPVCQNESSPFPTSTSSQSNPES---HCSDQPHDCSNDA- : 562
GmPRR59b1   : QSGIHPMSSPKPVCQNESSPFPTSTSTQSYPES---HNSDQLHDCSNDA- : 559
MtPRR59c    : QSGMHPTWTPKSVFQKESSPFPTSISSHSNPKS---QNSEP-HHWSDDAT : 459
CaPRR59c    : ES-------PKSVCQKESSSFPTSISSQSNPKN---QNSE----RSDDAT : 483
LcPRR59c    : QSGMNPTWTPKSMFQKESSPFPTSISSHSNPKS---QNSEA-RQWSDDTT : 463
Lcprr59c    : QSGMNPTWTPKSMFQKESSPFPTSISSHSNPKS---QNSEA-RQWSDDTT : 498
GmPRR59c2   : KSGVHPISTPKSVCQKESSPFPTSTSSQSNPQS---HNSER-HHWLEDAT : 544
GmPRR59c1   : ESGVHPIWTPKSVCQKESSPFPTITSSQSNPQS---HNSEC-HLWSEDST : 546
PvPRR59c    : QSVVHPIWTPKSVFQKESSPFPSSTSSQSNHQS---PNSKH-HHWSDDAT : 514
PtPRR59b1   : QS--NAAWNPNLAGRQQS-PFPTTASVHSNPEV---LDSKQNHKC----- : 548
PtPRR59b2   : QS--SAAWSPKLAGWQQS-PYPLSTSIHSNPDI---HDSEKNHRCSDETT : 575
OsPRR59     : GAIMQPMYYPQGAFMHCDSAAINKTAIQHVSCQSNYHENLGKPPQIDEHK : 557
BdPRR59     : GAIMQPIYYPQAPFMQHDPSAINQMAIQHASFHSNYHQSLGKPSEVVEHR : 575
```

```
                    *       820         *       840         *
OsPRR95     : RSHLSANRKHLRIEIPTDEPRHVSPTTGESGSSTVLDS-ARKTLSGSVCD : 545
BdPRR95     : LPAPNANEKHLHVEIPSDDPQHVSPMTGDSGSSTVLNN-SGNAPSGSGCD : 539
APRR5       : QDRDCSMDVDE--------RRYVSSATEHSAIG----------------- : 590
CaPRR59a    : NARNQIVYTQEHKSEHAEDQRLISPTTDQSVSSSFCN-GNASHLNSIGYG : 577
MtPRR59a    : STPNQIMYTQEHKSEHPEDQRLISPTTDQSVSSSFCNNGNASNFNSIGYG : 591
LcPRR59a    : SIPNQIVHTQEHRSEHAEDQRLISPTNDQSVSSSLCNNGNASHLNSIGYG : 580
GmPRR59a2   : TTQNHIVYTQEHKSENAEDQGHISPTTDQSVSSSFCN-GNASHLNSIGYG : 592
GmPRR59a1   : TTQNHIVYTQEHKSEHAEDRGHISPTTDQSVSSSFCN-GNASHLNSIGYG : 591
PvPRR59a    : SSQNHIVYTQEHKTEHAEVRGHISPTTDQSVSSSFCN-GNASHLNSIGYG : 582
PtPRR59a2   : DSINGSLQKQENKLDSLEGREHISSATDQSASSSFCN-GAASHFNSIGYG : 593
PtPRR59a1   : DSTNGSLQKQEDRLDSLEDRGLISPATDQSASSSFCN-GAASHFNSMGYG : 599
APRR9       : --VSFRN---QVLQSTVTNQKQDSPIPVESNREKAASK-EVEAGSQS--- : 387
MtPRR59b    : YSFHNQN---LTEKTDLDHAVHDSPSAGQGFGNDFCH--ASNHINSR--- : 533
LcPRR59b    : --FLNQN---VTEKYDLDHVVPDSPSRGPGFGNDICF--VSNHINS---- : 537
PvPRR59b    : -PCLNQN---VKEDTDLDQARHDSPAADQSTGNSLCHD-TSYHVNSSAYG : 574
GmPRR59b2   : -TCLDQN---VKDNTDSDHARHESPAADQSAGNNLCHD-AANHVNSSAYG : 607
GmPRR59b1   : -TCLNQN---VKDNTDSDHARHDSPVADQSAGNSLCHD-AANHVNSSAYG : 604
MtPRR59c    : Y-TCDQS---NND-----FAMHESPSNGQSC-TSFYHD-AESHNASGVCE : 498
CaPRR59c    : YSTCDQN---VNDQSTVDCAMHNSPASGQSSGTSFYHD-AVNHNASGVCE : 529
LcPRR59c    : Y-TCDRN---KNDRRNIDCAMHDSPSNGLSCGTSFYHD-AENRNTSGVCE : 508
Lcprr59c    : Y-TCDRN---KNDRRNIDCAMHDSPSNGLSCGTSFYHD-AENRNTSGVCE : 543
GmPRR59c2   : H-ASDQN---VNDQSNLECETHDSPAASQSAGPSFFHD-TANHNSSGVYR : 589
GmPRR59c1   : H-ASDKN---LNDQINLDCETHDSPDASQSAGTSFFHD-TANHNSSGVYR : 591
PvPRR59c    : H-ASDQN---VNDQSHLDFETQDSPTASQSADTALYHD-TTNHNCSGVYR : 559
PtPRR59b1   : --YVDQNDLQQNNREPVDEMRHDSPAAGQSTSSSLCNR-VANNNSSSAYE : 595
PtPRR59b2   : YNSVDQNDHQQNNKGPADEVRHDSPAAGQSTGG-LCNG-VINHNKSSAYE : 623
OsPRR59     : QPEENHQLHH-SRQILRESGEPVDLAKAHMERINQSASCSQDIRKGSGCT : 606
BdPRR59     : QLEENQLLHHHSRKILRES-EPIDLSRP--ENANPSTSCSQDLRRGSGWT : 622

                   860         *       880         *       900
OsPRR95     : ----SSSNHMIAPTE---SSNVVPENP---------------------- : 565
BdPRR95     : ----SSSNRIVAPLDPCNSFNGVPENPSM-------------------- : 564
APRR5       : ------NHIDQL--------IEKKNEDG--------------------YSL : 607
CaPRR59a    : SNCGSNSNVEQVATFRTATVSEGKNEEL-------------------TNS : 608
MtPRR59a    : SNCGSSGNVEQVATFRTAAVSEGKNEEL-------------------TNS : 622
LcPRR59a    : SNCGSSSNVETVTAFRTSAVSDGKNEDL-------------------TNG : 611
GmPRR59a2   : SNCGSSSNVDQVNTVW--AASEGKHKDL-------------------TSN : 621
GmPRR59a1   : SNCGSSSNVDQVNTVW--AASEGKHEDL-------------------TNN : 620
PvPRR59a    : SNCGSISNVDQVTTVR--VASESKNEDL-------------------TNN : 611
PtPRR59a2   : SASGSYSNADQIATVS--AASESKNEEG-----------------VFTHN : 624
PtPRR59a1   : STSGSNGNVDQVAIVR--DASESKNEEG-----------------AFTH- : 629
APRR9       : ----TNEGIAGQSSSTEKPKEEES-------------------------- : 407
MtPRR59b    : ----GNVGEAISNAVTKNSRTSSDGRRYNHS----NYDYDCDDDD--YEF : 573
LcPRR59b    : ----GNDERSTSNMVTENSRSSSD-RRYQKD----YYDYDD-------EF : 571
PvPRR59b    : SMDSGNDGHATSAIVSKNNPEGFSDSV--------CHNYDG--------S : 608
GmPRR59b2   : SMDSGNDGHATSAIVSKNTSDGFSDSG--------CHNYDG--------F : 641
GmPRR59b1   : SMDSGNDGNATSAIVSKNAPDGFSDSG--------CHNYDG--------F : 638
MtPRR59c    : GLGSVSDGNAPSTIVGKNNLESSMNNDH----------HDG--------L : 530
CaPRR59c    : GTGSGSDGNAPS-VVGNNNFESSMNNDH----------YDG--------L : 560
LcPRR59c    : GTGSGSDGNAPSSLVGKNHLESSISNDH----------YDE--------L : 540
Lcprr59c    : GTGSGSDGNAPSSLVGKNHLESSISNDH----------YDE--------L : 575
GmPRR59c2   : -----SDGNATSAKVAKESHEIFIDSGQR--------SYDG--------F : 618
GmPRR59c1   : SMGCRSDGNATSAKVAKESHGSFIDSGHC--------SYDG--------F : 625
PvPRR59c    : SIGCRSEGNATSAKVAKDNNESFFDIGHL--------GYDG--------F : 593
PtPRR59b1   : SFGSGNDVNASSVGTAEKSMAQENLNNGG------NFNHDG--------F : 631
PtPRR59b2   : SFGSRDDGNAK-----EKAMAQDNLNDGD------NFNRDG--------F : 654
OsPRR59     : GSGETDANTNTVIALESGNESGVQNCSN---------------------- : 634
BdPRR59     : VSGETDMNTNTIIAMESGNDSGIQNFSN---------------------- : 650
```

```
                          *       920         *       940         *
OsPRR95     : --DGLRHLSQREAALNKFRLKRKDRCFEKKVRYQSRKLLAEQRPRVKGQF : 613
BdPRR95     : --DGTHHLSQREVALNKFRLKRKERCFEKKVRYQSRKLLAEQRPRVKGQF : 612
APRR5       : SVGKIQQSLQREAALTKFRMKRKDRCYEKKVRYESRKKLAEQRPRIKGQF : 657
CaPRR59a    : GYSH--RSILREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 656
MtPRR59a    : GYSH--RAMLREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 670
LcPRR59a    : GYSHSHRSMLREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 661
GmPRR59a2   : ANSH--RSIQREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 669
GmPRR59a1   : ANSH--RSIQREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 668
PvPRR59a    : ANSH--RSIQREAALNKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 659
PtPRR59a2   : SNSH--RSIQREAALTKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 672
PtPRR59a1   : SYSH--RSIQREAALTKFRLKRKERCYEKKVRYESRKKLAEQRPRVKGQF : 677
APRR9       : -AKQRWSRSQREAALMKFRLKRKDRCFDKKVRYQSRKKLAEQRPRVKGQF : 456
MtPRR59b    : RLSDSHRSRQREAALTKFRLKRKERCFEKKVRYQSRKKQAEQRLRVKGQF : 623
LcPRR59b    : RSSDSHRSSQREAALTKFRLKRKERCYAKKVRYQSRKRLAEQRLRVKGKF : 621
PvPRR59b    : RATYSHRTSQREAALTKFRLKRKERCFEKKVRYQSRKRLAEQRPRVKGQF : 658
GmPRR59b2   : RVTDSHRSSQREAALVKFRLKRKERCFEKKVRYQSRKRLAEQRPRVKGQF : 691
GmPRR59b1   : RVTDPHRSSQREAVLVKFRLKRKERCFEKKVRYQSRKRLAEQRPRVKGQF : 688
MtPRR59c    : RDTSSHRTSQREAALTKFRLKRKDRCYDKKVRYESRKRQAENRPRVKGQF : 580
CaPRR59c    : RGTNSHRTSQREAALTKFRLKRKDRCYEKKVRYESRKRLADNRPRVKGQF : 610
LcPRR59c    : RGPNSHHTSQREAALTKFRLKRKERCYDKKVRYESRKRQADKRPRVKGQF : 590
Lcprr59c    : RGPNSHHTSQREAALTKFRLKRKERCYDKKVRYESRKRQADKRPRVKGAV : 625
GmPRR59c2   : IGTDSHRTSQREAALTKFRLKRKDRCYEKKVRYQSRKRLAEQRPRVKGQF : 668
GmPRR59c1   : IGTDSHRTSHREAALTKFRLKRKDRCYEKKVRYQSRKRLAEQRPRVKGQF : 675
PvPRR59c    : IGTDSHRTSQREAALTKFRLKRKDRCYEKKVRYQSRKRLAEQRPRVKGQF : 643
PtPRR59b1   : GGSDSYRSSQREAALTKFRLKRKDRCYEKRVRYQSRKRLAEQRPRVKGQF : 681
PtPRR59b2   : RGIDSLRSSQREAALTKFRLKRKDRCYEKKVRYQSRKRLAEQRPRVKGQF : 704
OsPRR59     : -NVLDGDRSRREAALLKFRMKRKDRCFEKKVRYHSRKKLAEQRPRVKGQF : 683
BdPRR59     : -NGLDIDRSRREAALMKFRMKRKDRCYEKKVRYHSRKKLAEQRPRIKGQF : 699

                        960         *       980
OsPRR95     : VRQDHGV---QGS-------------------- : 623
BdPRR95     : VRQEPSI---QGS-------------------- : 622
APRR5       : VRQVQST---QAP-------------------- : 667
CaPRR59a    : VRQVNPE-SIAAEKDGNEY------------- : 674
MtPRR59a    : VRQPNPD-SLSGEKDC---------------- : 685
LcPRR59a    : VRQVNPD-SLSGEKDC---------------- : 676
GmPRR59a2   : VRQVHPD-PLVAEKDCKEYDHSHISDTLERRA : 700
GmPRR59a1   : VRQVHPD-PLVAEKDGKEYDHSDF-------- : 691
PvPRR59a    : VRQVLPD-GLVAEQDGKEYDHSQISDLLERRA : 690
PtPRR59a2   : VRQVHID-PSPAETDQ---------------- : 687
PtPRR59a1   : VRQVHID-PSPAETDQ---------------- : 692
APRR9       : VRTVNSD--ASTKS------------------ : 468
MtPRR59b    : VRKVQNDDHPNVDS-GDQ-------------- : 640
LcPRR59b    : IHRVHDDDHPNADANGDQ-------------- : 639
PvPRR59b    : VRQVHND-HPVADAGGDS-------------- : 675
GmPRR59b2   : VRQHN---HPFAEAGGDS-------------- : 706
GmPRR59b1   : VRQHD---HPVAEAGGDS-------------- : 703
MtPRR59c    : VRQVQG-EVPVAETRGY---------------- : 596
CaPRR59c    : VRQVHS-EHPAADAGGY---------------- : 626
LcPRR59c    : VRQVQS-ELAVDIG-GY---------------- : 605
Lcprr59c    : CAPSTK--------------------------- : 631
GmPRR59c2   : VRQVHD-DHPVADVGGGS-------------- : 685
GmPRR59c1   : VRQVQD-DHPVADVGGGS-------------- : 692
PvPRR59c    : VRQVHS-DHPVSDLGGDS-------------- : 660
PtPRR59b1   : VRQAQN-DCPVANG------------------ : 694
PtPRR59b2   : VRQVQN-DSPIANG------------------ : 717
OsPRR59     : VSQKLKS-AITT-----EAETD---------- : 699
BdPRR59     : VSQKLKS-DTATPTTTEDVETD---------- : 720
```

The alignment was created with full-length protein sequences of *L. culinaris* (Lc), *M. truncatula* (Mt), *C. arietinum* (Ca), *G. max* (Gm), *P. vulgaris* (Pv), *A. thaliana* (At). *O. sativa* (Os), *B. distachyon* (Bd), and *P.* trichocarpa (Pt) aligned with ClustalX and manually adjusted and annotated using GeneDoc and Adobe Illustrator. Shading indicates degrees of conservation; black=100%, dark grey=80%, light grey=60%, yellow=frame-shift. Lcprr59c is the LcPRR59c predicted protein from ILL 2601. Refer to Appendix 4 for sequence information.

Appendix 9 *Lentil linkage group nomenclature*

Medicago (Mt4.0)	This study (Chapter 4)	Kaur et al. (2014)	Sharpe et al. (2013)
1	1, 5	3, 5.1	1, 5
2	2	4.1, 4.2	2
3	3	3, 6	3
4*	4, 7*	1, 2.1, 2,2*	4*
5	1, 5	5.1, 5.2	1, 5
6	2	4.1	2
7	6	7	6
8*	4, 7*	1, 2.2*	7*

Asterisk (*) refer to translocations attributable to an aberrant chromosomal arrangement that resulted from a translocation event in the *M. truncatula* model accession A17, where the reciprocal translocation of the long arms of chromosomes 4 and 8 has been noted (Kamphuis et al. 2007). It is likely Sharpe et al. (2013) used older Mt3.0 genome assembly (no translocation) to establish relationship between *M. truncatula* chromosomes and presented lentil linkage groups.